

Gly Gly Asg Asg Xaa Phe  
20

## (2) INFORMATION FOR SEQ ID NO:130:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 7 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Asp Pro Gly Tyr Thr Pro Gly  
1 5

## (2) INFORMATION FOR SEQ ID NO:131:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ix) FEATURE:

(D) OTHER INFORMATION: /note= "The Second Residue Can Be Either a Pro or Thr"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Xaa Xaa Gly Thr Thr Gly Pro Gln Phe Tyr  
1 5 10

## (2) INFORMATION FOR SEQ ID NO:132:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 9 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (ix) FEATURE:

(D) OTHER INFORMATION: /note= "The Third Residue Can Be Either a Val or Leu"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Xaa Pro Xaa Val Thr Ala Tyr Ala Gly  
1 5

## (2) INFORMATION FOR SEQ ID NO:133:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:133:

Xaa Xaa Xaa Glu Lys Pro Phe Leu Arg  
1 5

(2) INFORMATION FOR SEQ ID NO:134:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:134:

Xaa Asp Ser Glu Lys Ser Ala Thr Ala Lys Val Thr Asp Ala Ser  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:135:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Ala Gly Asp Thr Xaa Ile Tyr Ile Val Gly Asp Leu Thr Ala Asp  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:136:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Ala Pro Glu Ser Gly Ala Gly Leu Gly Gly Thr Val Glu Ala Gly  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

Xaa Tyr Ile Ala Tyr Xaa Thr Thr Ala Gly Ile Val Pro Gly Lys Ile  
1 5 10 15  
Asn Val His Leu Val  
20

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 892 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

GCAAGCTGT	CGTGGCTTT	GGGTGATG	GTTCACCTC	GCTAGCGTG	GGGTGCGCG	60
TCACCATCG	ACGACCGCG	GCCTCAAAAC	CGGTAGAGG	ACACCAAAAC	GCCTAGCGCG	120
GGAGTTCAT	GGCTTTTTC	CGACGCAAC	AGCAGGCGC	GCTCGCGCG	GCTCGCGCG	180
ATGATCCAC	CGCTGATTC	CAGGCGCGA	CGATTCGCG	TGTACAGAG	GTGTGCGCG	240
GGTGGTAC	CTACCGCG	GTGGTGGGA	GGCGGCTTC	GCTGCGCGC	GAGGCGCGC	300
CGTGGCGCG	TGTGTGCGT	GGCGGCGCG	CAATCGCGT	CGCATCATC	ATTGCGCGT	360
TCGCGGCTT	GGCGCGTGA	ATCGCGCGA	TGCGCGCGC	ACCGCGCGA	ACCGCGCGA	420
CGAGTGGCG	GAGCGCGCG	CGACCGCGC	CGCGCGCGC	GGCGCGCGC	ACCGCGCGA	480
GGAGCGCGC	GAGCGCGCG	GTGCGCGCG	CGCGCGCGC	GCGCGCGCG	ACCGCGCGA	540
CGAGCGCGC	AGCGCGCGC	CGCGCGCGA	CGCGCGCGC	GCGCGCGCG	GCTCGCGCG	600
CGTGGCGCG	GAGCGCGCG	GTCGCGCGA	CGCGCGCGC	GAGCGCGCG	GCTCGCGCG	660

GGAGGACAGCA GCGGACGGCA GACCAAGCC AACAGATGCC AACCCAGAG CAGACCGTGG	726
CCGCGCAGAG GGTGGGCGCC GCTGGCGAGC CCGCGTCGG TGGCGGAGC GGCAGCGCC	780
GGGCGACTT ATTGGCGGG TTCTGATCAC GCTCGCGGT TCACTACGT CCGAGGACAT	840
GCGCGTGTAT GCGTGACGG TCGTCTGAC CTCTCTCAG CA	892

## (2) INFORMATION FOR SEQ ID NO:139:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 815 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

## (1) SEQUENCE DESCRIPTION: SEQ ID NO:139:

CGATCAACCA ACGGTGCGG CGGCGCGCG CCGCGGATCC GCGTGGCGG GCACGCGCGC	80
CGGTGCTTTC GTTGGCGCGG TTGCGCGGT CCGCGCGGTC GCGCGCGAGC GGTGGGTTC	120
GTAGGCGGCT GTTACGCGC TGGTTGGCG GGCAGCGGCT GGCACCGCG GTACGCGCGA	180
TGGCGCGGTT GCGCGCGCG GGCAGCGGTC CAGCGGTGCG ACGGTGCGA CGGTGCGCGA	240
CGACCGCGCG GCGCGCGCG CCGCGCGCG CCGCGCGCG CCGCGCGCG GGTGCGCGGT	300
TCTGCGCGTT ACGCGCGCG CCGCGCGTC CCGCGCGCG CCGCGCGCG GTACGCGCGG	360
ACCGCGCGCG CCGCGCGCG CCGCGCGCG CCGCGCGCG ACGCGCGCG CCGCGCGCG	420
GGAGTGGCGG GTTGGCGCG GTACGCGCG GGCAGCGCG AACTACTTTC GGTGCGCGAG	480
CAGTTCCAGA CGACCGCGA GCGCGCGGT GTGCGCGCG TGGTGAAT GCGCGCGAT	540
AGCGCGTAC TGTGCGTTC GGTGCGGTC GATGATGTC TCGAGGTAC GGTGCGCGG	600
CGCGCGGAG GAGCGCGTGA ACTGCGGTT GCGCGGATCG GCGGTGCTT GCGCGGATTC	660
CGAGCGCAT ACGCGGATC CGGTGCGTA ACGCGCGCG ACGCGGCTT GGTGCGCGG	720
AGCTGCTGCG GGTGCGGTC TTACGCGGTT GTGCGGAG ACGAGTACCA GGTGCGTTC	780
GCGCGCGCGA TCGACCGCG GTGCGGTC GTCT	815

## (2) INFORMATION FOR SEQ ID NO:140:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1152 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(11) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

ACGAGCCGCC GCGTGAGGTC TCGATCAGA GAGTCTCGG ACTCAGCGG GCGGTTGACC	60
CTCTCCCAAG AACAACTGCT GAGATCTCT GCGCGCGAAA CAGCGCGTGA TTGAGCGCTC	120
TATGACGGGT TGACAGAGGA GATGATCGG CAGATTGATA TGGACCGGCT GGGCTAAGAG	180
GTGAGCAAGA TGGTGCAGCT ATATGCTCG GACTCCGCT CCGGATCAG CTTCGCGAC	240
GCGCGGCTGA TGGTGGAGAG CGAGAGCTC GCGAGAGCC AGTATCGAT CGAGAGCGTC	300
GACGGCATCA CGCTGTTGG GCGGCGGAG ATGACAACGC CTTTCATGCT TGAGATGCTC	360
AAGCGTGACC GCGACTTCCA GCTCTTCAG ACCAGCGGC ACTACAGGG CGCGATTCCA	420
ACGCGCGAGG TGTGATACG GCGCGGCTC GGTCAAGCAG TTCACTGAG CAGCGATGCT	480
GCGTTCTGCT TGCTGTTAG CAACCGGTC GTCTGAGGA AGATGTTGAA TGACGAGGCC	540
TTGATTGCGG CACACAGCTC GGGCAAGAC CTTCGTGAG GATCTGGAC GATGAGGAC	600
TGCTGGGCT GCGTGCATCG ATCGGCTCC CTGCGGAGT TGAAGCGCT CAGGCGAAAT	660
GCGCAAAAGG CATACTTCAC GCGCTGGA CATCTGCTC GCGAGGATC CGCATTCAG	720
GCGGCTTCCA CTGCGCGCC GTTGAAGCC TTCACTGGA TGCTGAGCT CGGCTATTGC	780
CTGCTGTACA AGAATATCT AGGCGGATC GAGCTGCA GCGTGAAGGC GTATATGCT	840
TGCTATACAT AGGATTACG AGGCGATGCA AGTCTGCTC CGGATTCGG CAGGAGCTCC	900
GCTGAACCGG CTGCGCGCT GTCATGCTC CATTGTAAT CGCTGCTCC CAGGCGGACC	960
CGCGGCGGGA ATACGACAG ATCGGACAG GATTGCGGC GAGCGGCTT GAGGCGCTGC	1020
ATCGCGCGG CACATCTAC GCGAGGAAAC AGCGCTGCA CGTGGCGCT GCGGCTGTCC	1080
GCGTACTTT CAGACCGGCC CATCACTTGA TCACAGCTCG GCGGACTTC CATTGCTGTC	1140
GTTCGGCAGC AG	1152

## (2) INFORMATION FOR SEQ ID NO:141:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 655 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

CTGCTGCCAA TTGCGAGGG TGTACTGCG GTTGAGTAT GCTGCTGAG TCGCGACGAC	60
CAGCAATGCG GCAACGCGAC GATTCGCGT CAAGGAGGCC AGCGGCTCCA CCGCGCGGAT	120

```

CGCTCGAGT CGCGCTGGG CGCTCTTTC CTGGGGGAG GTCATCGAG GTTTTCGGC 180
CTTGGTTTC CGCATTAAG CGGCGCGCG GGTGGGGCG GCGGTATGG CGGAAGTGG 240
ATCAGCACAC CGAGATAGG GGTCTGTGA AGCTTTTGA GGTGCGGCG GGGCACTTC 300
CGCGCAAT CTACTAGCA GAAGTCTGG CGCATACGA TCTGACCAA CTGGTGGGG 360
TGCAGCCAC CTTCTATGG GATGGCGCG AGATGGCGG CTGACCAAT CTATGCGCG 420
TTGGGACG CGACCGATA GGTGTCAG TGGGTCTAC GCTTGGGCT TTGGGACGG 480
TGGGAGCT GGTGGGCTT GCGCGCGAA AGCGGGGGT CGGTGGCAT CAGGAATCG 540
TCAAGCGCG GCGACTGCG CGCGCTGCG GCGGGATGT CAGCATCGG GATCTATCG 600
TGGGTTCAT ACTCTGAG CACTGGGCG AACAGTCA TTGGGAGC GCGCA 660

```

## (i) INFORMATION FOR SEQ ID NO:142:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 567 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) SEQUENCE DESCRIPTION: SEQ ID NO:142:

```

Asn Ala Val Val Ala Phe Ala Val Ile Gly Phe Ala Ser Leu Ala Val 1
5
Ala Val Ala Val Thr Ile Arg Pro Thr Ala Ala Ser Lys Pro Val Glu 20
25
Gly His Gln Asn Ala Gln Pro Gly Lys Phe Met Pro Leu Leu Pro Thr 35
40
Gln Gln Gln Ala Pro Val Pro Pro Pro Pro Asp Asp Pro Thr Ala 50
55
Gly Phe Gln Gly Gly Thr Ile Pro Ala Val Gln Asn Val Val Pro Arg 65
70
Pro Gly Thr Ser Pro Gly Val Gly Gly Thr Pro Ala Ser Pro Ala Pro 85
90
Gln Ala Pro Ala Val Pro Gly Val Val Pro Ala Pro Val Pro Ile Pro 100
105
Val Pro Ile Ile Ile Pro Pro Phe Pro Gly Trp Gln Pro Gly Met Pro 115
120
Thr Ile Pro Thr Ala Pro Pro Thr Thr Pro Val Thr Thr Ser Ala Thr 130
135
Thr Pro Pro Thr Thr Pro Pro Thr Thr Pro Val Thr Thr Pro Pro Thr 140

```

150

```

145              150              155              160
Thr Pro Pro Thr Thr Pro Val Thr Thr Pro Pro Thr Thr Pro Pro Thr
              165              170              175
Thr Pro Val Thr Thr Pro Pro Thr Thr Val Ala Pro Thr Thr Val Ala
              180              185              190
Pro Thr Thr Val Ala Pro Thr Thr Val Ala Pro Thr Thr Val Ala Pro
              195              200              205
Ala Thr Ala Thr Pro Thr Thr Val Ala Pro Gln Pro Thr Gln Gln Pro
              210              215              220
Thr Gln Gln Pro Thr Gln Gln Met Pro Thr Gln Gln Gln Thr Val Ala
              225              230              235              240
Pro Gln Thr Val Ala Pro Ala Pro Gln Pro Pro Ser Gly Gly Arg Asn
              245              250              255
Gly Ser Gly Gly Gly Asp Leu Phe Gly Gly Phe
              260              265

```

## (2) INFORMATION FOR SEQ ID NO:143:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 174 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (iii) SEQUENCE DESCRIPTION: SEQ ID NO:143:

```

Ile Asn Gln Pro Leu Ala Pro Pro Ala Pro Pro Asp Pro Pro Ser Pro
1      5      10      15
Pro Arg Pro Pro Val Pro Pro Val Pro Pro Leu Pro Pro Ser Pro Pro
20     25     30
Ser Pro Pro Thr Gly Trp Val Pro Arg Ala Leu Leu Pro Pro Trp Leu
35     40     45
Ala Gly Thr Pro Pro Ala Pro Pro Val Pro Pro Met Ala Pro Leu Pro
50     55     60
Pro Ala Ala Pro Leu Pro Pro Leu Pro Pro Leu Pro Pro Thr
65     70     75     80
Ser His Pro Pro Arg Pro Pro His Pro Pro Ala Pro Pro Ala Pro Pro
85     90     95
Ala Cys Pro Phe Val Pro Val Pro Pro Ala Pro Pro Leu Pro Pro Ser
100    105    110
Pro Pro Thr Gln Leu Pro Ala Asp Ala Ala Cys Pro Pro Ala Pro Pro
115    120    125

```

151

Ala Pro Pro Leu Ala Pro Pro Ser Pro Pro Ala Gly Ser Ala Ala Ile  
 130 135 140  
 Arg Ala Leu Thr Gly Ala Thr Ser Ala Ser Thr Leu Gly His Arg Ala  
 145 150 155 160  
 Leu Pro Asp Asp Thr Thr Ala Arg Gly Cys Arg Arg Thr Gly  
 165 170

## (2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

Gln Pro Pro Ala Glu Val Ser Asp Gln Arg Val Ser Gly Leu Thr Gly  
 1 5 10  
 Ala Val Gln Pro Ser Pro Arg Thr Thr Ala Glu Asp Pro Arg Pro Arg  
 20 30  
 Asn Arg Arg  
 35

## (2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 104 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

Arg Ala Arg Ser Ala Gly Cys Thr Cys Arg Trp Cys Asn Pro His Glu  
 1 5 10  
 Cys Arg Arg Pro Ala Met Arg Gln Gln His Gly Ser Arg Ser Thr Thr  
 20 25 30  
 Pro Pro Gly Pro Arg Gly Arg Ser Ala Arg Val Arg Pro Gly Arg Leu  
 35 40 45  
 Phe Pro Trp Ala Gly Ser Ser Asp Val Phe Pro Ser Trp Phe Ala Ala  
 50 55 60  
 Ile Met Pro Ala Arg Arg Val Gly Arg Pro Val Trp Pro Asn Val Asp  
 65 70 75 80



Gln His Thr Arg Asp Thr Gly Leu Cys Lys Leu Phe Glu Arg Arg Ala  
 85 90 95

Gly Gln Leu Arg Arg Gln Phe Tyr  
 100

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

GGATCATAT GGGCATCAT CATCATCATC AAGTGATGGA CATCATGGG AGC

53

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR Primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

CTGTGAATGCA GAGCTCTGGGT GGGCGGCGCT CATCTTGAAC GA

42

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR Primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:148:

GGATCTGCA GGCTCGAAAC CACCGAGGG T

31

(2) INFORMATION FOR SEQ ID NO:149:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(11) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PCR primer"

(15) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CCTGGAATTG AGCGCTGGAATTCCTCGGA T

31

(2) INFORMATION FOR SEQ ID NO:150:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(11) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PCR primer"

(15) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:150:

GGATCCAGCG CTCGATCGAA GACCATGCG CTT

33

(2) INFORMATION FOR SEQ ID NO:151:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(11) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PCR primer"

(15) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:151:

GAGAGAAATC TCAGAAAGCC ATTTCGAGG ACA

33

(C) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1993 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycobacterium tuberculosis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 152..1273

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

TGTCCTCCGA	CGGCGGGCTG	GTGGAGGAGG	GGCCACCGGA	ACAGCTCTTC	TCCTCGCCGA	60
AGCTGTCGGA	AACCGCCCA	TACCTGCGCG	GGCTGTGCGG	GGAGCTCAAG	GAGGCCAAGC	120
GGCGAATTC	AAGAGACAG	AAAGGTATGG	C	CTG	AAA	172
				Val	Lys	
					Arg	
					Leu	
					His	
					Thr	
CPG	TTG	GCC	CTG	TTG	ACC	220
Leu	Leu	Ala	Val	Leu	Thr	
					Ala	
					Ala	
					Pro	
					Leu	
					Leu	
					Ala	
					Ala	
					Gly	
TTT	GCC	TGG	AAA	GCA	CGG	260
Cys	Gly	Ser	Lys	Pro	Ser	
					Gly	
					Ser	
					Pro	
					Glu	
					Thr	
					Gly	
					Ala	
					Gly	
					Ala	
GGT	ACT	GTC	GCG	ACT	ACC	316
Gly	Thr	Val	Ala	Thr	Thr	
					Pro	
					Ala	
					Ser	
					Ser	
					Pro	
					Val	
					Thr	
					Leu	
					Ala	
					Glu	
ACC	GGT	ACC	ACC	CTG	CTC	364
Thr	Gly	Ser	Thr	Leu	Tyr	
					Pro	
					Leu	
					Pro	
					Asn	
					Leu	
					Thr	
					Gly	
					Pro	
					Ala	
TTT	CAG	GAG	ACG	TAT	CGG	412
Phe	His	Glu	Arg	Tyr	Pro	
					Asn	
					Val	
					Thr	
					Ile	
					Thr	
					Ala	
					Gln	
					Gly	
					Thr	
					Gly	
TTT	GAT	GCC	CGG	ATC	GCG	460
Ser	Gly	Ala	Gly	Ile	Ala	
					Gln	
					Ala	
					Ala	
					Gly	
					Thr	
					Val	
					Asn	
					Ile	
					Gly	
GCC	TCC	GAC	GCG	TAT	CTG	508
Ala	Ser	Asp	Ala	Tyr	Leu	
					Ser	
					Gln	
					Gly	
					Asp	
					Met	
					Ala	
					Ala	
					His	
					Lys	
					Gly	
CTG	ATG	AAC	ATC	GCG	CTA	556
Leu	Met	Asn	Ile	Ala	Leu	
					Ala	
					Ile	
					Ser	
					Ala	
					Gln	
					Gln	
					Val	
					Asn	
					Tyr	
					Asn	

CTG CCG GGA GTG AGC GAG CAG CTC AAG CTG AAC GGA AAA GTC CTG GCG Leu Pro Gly Val Ser Glu His Leu Lys Leu Asn Gly Lys Val Leu Ala 140 145 150	584
GCG ATG TAC CAG GGC ACC ATC AAA ACC TGG GAG GAC CCG CAG ATC GGT Ala Met Tyr Gln Gly Thr Ile Lys Thr Trp Asp Asp Pro Gln Ile Ala 155 160 165	652
GCG CTC AAC CCG GCG GTG AAC CTG CCG GCG ACC GCG GTA GPT CCG CTG Ala Leu Asn Pro Gly Val Asn Leu Pro Gly Thr Ala Val Val Pro Leu 170 175 180	700
CAC CAG TCC GAC GGG TCC GGT GAC ACC TTC TTG TTC ACC CAG TAC CTC His Arg Ser Asp Gly Ser Gly Asp Thr Phe Leu Phe Thr Gln Tyr Leu 185 190 195	748
TTC AAG CAA GAT CCG GAG GCG TCG GCG AAG TCG CCG GCG TTC GCG ACC Ser Lys Gln Asp Pro Glu Gly Trp Gly Lys Ser Pro Gly Phe Gly Thr 200 205 210 215	796
ACC CTC GAC TTC CCG GCG GTG CCG GGT GCG CTG GGT GAG AAC GCG AAC Thr Val Asp Phe Pro Ala Val Pro Gly Ala Leu Gly Glu Asn Gly Asn 220 225 230	844
GCG GCG ATG CTG ACC GGT TCG GCG GAG ACG CCG GCG TCG CTG GCG TAT Gly Gly Met Val Thr Gly Cys Ala Glu Thr Pro Gly Cys Val Ala Tyr 235 240 245	892
ATC GCG ATC AGC TTC CTC GAC CAG CCG AAT CAA CCG GGA CTC GCG GAG Ile Gly Ile Ser Phe Leu Asp Glu Ala Ser Gln Arg Gly Leu Gly Glu 250 255 260	940
GCG CAA CTA GCG AAT AGC TCT GCG AAT TTC TCG TTG CCG GAG GCG GAA Ala Gln Leu Gly Asn Ser Ser Gly Asn Phe Leu Leu Pro Asp Ala Glu 265 270 275	988
AGC ATT CAG GCT CCG GCG GGT GCG TTC GCA TCG AAA ACC CCG GCG AAC Ser Ile Gln Ala Ala Ala Ala Gly Phe Ala Ser Lys Thr Pro Ala Asn 280 285 290 295	1036
CAG GCG ATT TCG ATG ATC GAC GCG CCC GCG CCG GAC GCG TAC CCG ATC Gln Ala Ile Ser Met Ile Asp Gly Pro Ala Pro Asp Gly Tyr Pro Ile 300 305 310	1084
ATC AAC TAC GAG TAC GCG ATC CTC AAC AAC CCG CAA AAG GAC GCG GCG Ile Asn Tyr Glu Tyr Ala Ile Val Asn Asn Arg Gln Lys Asp Ala Ala 315 320 325	1132
ACC GCG CAG ACC TTC CAG GCA TTT CTG CAC TCG GCG ATC ACC GAC GCG Thr Ala Gln Thr Leu Gln Ala Phe Leu His Ala Trp Ala Ile Thr Asp Gly 330 335 340	1180
AAC AAG GCG TCG TTC CTC GAC CAG GPT CAT TTC CAG CCG CTG CCG CCG Asn Lys Ala Ser Phe Leu Asp Gln Val His Phe Glu Pro Leu Pro Pro 345 350 355	1228
GCG GTG HTG AAG TTG TGT GAC GCG TTG ATC GCG ACC ATT TCG AGC Ala Val Val Lys Leu Ser Asp Ala Leu Ile Ala Thr Ile Ser Ser 360 365 370	1276
TAGCGCTGTT AACACACACG CGACACACAC CTGCGTGGCG CCATCGGGGT GCTTTCGGGA	1332

```

GGATGCTGGC CGGTGCGGGT GAAGTGGGCT GGGCTGAGCC GCGCATCGGG TGGTTGGGTC 1393
GGATAGGTGC GTGATGCGCG CTGCTTGGCG TGGTCTTGGT GTTGGTGGTC CTGCTATGCG 1403
AGCGCATGGC TGGGATCAGG CTCAAGGGGT TGGATTCTT GACCGGACCG GAATGGAAATC 1513
CAGGCAACAC CTACGGCGAA ACGGTGTGCA CGAGCGGCTC GCGCATCGGG TGGGCGCTTA 1573
CTACGGGGCG TTGCGCTGCA TCGTGGGAG GCTGGGGAGC TGGGTAATCG CGCTGATGAT 1633
CGCGGTGCTG GTCTCTGTAG GAGCGCGGCT GGTGATCTTC GAACGGCTCG CGAAACGGTT 1693
GGCGGAAGAT GTGGCAATAG TCGTGGAATT GCTGCGCGGA ATCGCGAGCG TGGTCTTGGC 1753
TTTGTGGGGG GGAATGACGT TCGCGCGCTT LBTGCTCAT GACATCGCTC CGGTGATCGC 1813
TGACAACGGT CCGGATGTCG CGGTGCTGAA CTACTTGCGC GCGGACGCGG GCAACGGGGA 1873
GCGCATCTTG GTTCTGGGTC TGGTGTGGC GTTGATCTTC GTTGCATTA TGGCAGCCAC 1933
CACTCATGAC CTGTTCCGGC AGGTGCGGCT GTTGGCGCGG GAGGCGGGA TGGGGAATTC 1993

```

(2) INFORMATION FOR SEQ 10 NO:153:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) SEQUENCE DESCRIPTION: SEQ 10 NO:153:

```

Val Lys Ile Arg Leu His Thr Leu Leu Ala Val Leu Thr Ala Ala Pro
 3           5           10           15
Leu Leu Leu Ala Ala Ala Gly Cys Gly Ser Lys Pro Pro Ser Gly Ser
25           30
Pro Gln Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro His Ser
35           40           45
Ser Pro Val Thr Leu Ala Ala Thr Gly Ser Thr Leu Leu Tyr Pro Leu
50           55           60
Phe Asn Leu Arg Gly Pro Ala Phe His Glu Arg Tyr Pro Asn Val Thr
65           70           75           80
Ile Thr Ala Gln Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln Ala Ala
85           90           95
Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser Gln Gly
100          105          110
Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala Ile Ser
115          120          125
Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His Leu Lys
130          135          140

```

157

```

Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Glu Gly Thr Ile Lys Thr
145          150          155          160

Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn Leu Pro
165          170          175

Gly Thr Ala Val Val Pro Leu Ala Arg Ser Asp Gly Ser Gly Asp Thr
180          185          190

Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Gln Gly Trp Gly
195          200          205

Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val Pro Gly
210          215          220

Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys Ala Glu
225          230          235          240

Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala
245          250          255

Ser Gln Arg Gly Leu Gly Glu Ala Glu Leu Gly Asn Ser Ser Gly Asn
260          265          270

Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Gly Phe
275          280          285

Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro
290          295          300

Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile Val Asn
305          310          315          320

Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu
325          330          335

Ala Tip Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val
340          345          350

His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu
355          360          365

Ile Ala Thr Ile Ser Ser
370

```

## (2) INFORMATION FOR SEQ ID NO:154:

- (a) SEQUENCE CHARACTERISTICS:
- (i) LENGTH: 1993 base pairs
  - (ii) TYPE: nucleic acid
  - (iii) STRANDEDNESS: single
  - (iv) TOPOLOGY: linear

## (a1) SEQUENCE DESCRIPTION: SEQ ID NO:154:

TTTCTTCGTA CCGCAGGAGTGA GTGAGGAGGAG GAGCCAGCCCA ACAGCTGTTC TCCTGGCGGA

AGCATGCGGA	AGCCGCCGGA	TACTTCGGGG	GACTGTCTGG	GGACGTCAAG	GACGCCAAGC	120
GCGGAAGTTG	AAGAGACGCG	AAAGGTATGG	CGTGAAGATT	CGTTTCGATA	CGCTGTTGGC	180
CGTTTTCAGC	GCTGGCGTGC	TGTTGCTAGC	AGCGGCGGGC	TGTGTCTCGA	AAGCAGCGAG	240
CGTTTCGGCT	GAAGCGGGCG	CGGGCGCGGG	TACTGTGCGG	ACTACCGCGG	CGTGTCTGGC	300
GCTGAGCTTG	CGGAGACCGG	GTAGCAGCTT	GCTGTACCGG	CTGTTCAAGC	TGTGGGTTTC	360
GGCTTTTCAC	GAGAGGTATC	CGAAGGTGAC	GATCACCGCT	CAGGSCACCG	GTCTGTGCTC	420
CGGATTCGGC	CAGGTGCGCG	CGGGACGGGT	CAACATTGCG	GGCTCGAGCG	CGTATCTGTC	480
GGAGGTGAT	ATGGCGCGCG	NCAAGGGGCT	GATGAACATC	GGCTGAGCCA	TCTCGGCTCA	540
GGAGGTCAAC	TACAGCTGCG	CGGAGGTGAG	CGAGCACCTC	AAGCTGAAGC	GAAAGGTGCT	600
GGCGCGGATG	TACGAGGCGC	CGATCAAAAC	CTGGGACGAC	CGCGGAGTCC	CTGCGCTCAA	660
CGCGCGGCTG	AACCTGCGCG	GCACCGCGGT	AGTTCGGCTG	CACCGCTGCG	AGGGCTCGCG	720
TGACAGCTTC	TCTTTCAGCT	AGTACCTGTC	CAGCAGGAT	CTCGAGGGCT	GGGTCAAGTC	780
CGCGGCTTC	CGGACGAGCG	TGAGCTTCCC	GGGGTGTGCG	GGTTCGCTCG	GTGAGACCGG	840
CAGCGCGCGC	ATGCTGAGCG	GTTCGCGCGA	GACACCGCGC	TGCTGGGCTT	ATATCGGCTT	900
CAGCTTCCTC	GAGCAGGCGC	GTGAGGTGCG	ACTGCGGAG	CGCGAGGTAG	GGATAGCTTC	960
TGGGAATTTC	TGTGTGCGCG	AGCGCGAAGG	CATTGAGGCG	CGCGGCGCTG	GGTTGCTGTC	1020
GAAAGCCCGC	CGGACGAGCG	CGATTTCTAT	GATCGAGGCG	CTCGCGCTCG	ACGGGTAGCT	1080
GATCATGAC	TAGGATGAG	CGATGCTGAA	CAGCGGCAAG	AAGGAGCGCG	CGACCGCGCA	1140
GACCTTCGAC	CGATTTCTCG	ACTGGGCGAT	CAGCGACCGG	AGCAGGCGCT	CGTTCTTCTG	1200
CGAGTTCTAT	TTCGAGCGCG	TGCGCGCGCG	GGTGTGAGCG	TGTGTGAGCG	CGTTGATGCG	1260
GAGCATTTCC	AGCTAGCTTC	GTGAGCAGCG	ACGCGACAGC	AGCTCTGCTC	GGGCACTGCG	1320
CGCTGTTTGG	GGGATCTGCT	GGCGCGTGGC	GGTGAATGCG	CGCGCGCTCG	CGCGGCGATC	1380
CGGTGCTTGG	GTGGATGAGC	TGCGGTGATC	CGGCTGCTTG	CGCTGCTGCT	GGTGTGCTGCG	1440
GTGCTGCTGA	TGGAATGCTAT	GGTGTGCTAT	AGCTTGAAGC	CGTTGCTATTT	GTTCAGCGCG	1500
AGCGAATGCA	ATCGAGGCGA	CAGCTAGCGC	GAAGCGCTTG	TGCGCGGCGC	GTGCGCGATC	1560
CGGTGGCGCG	GTATAGCGCG	GGTTTGGCGC	TGATCGCTCG	GAGCGTGGCG	AGTTCGCGCA	1620
TGCGCTGCTAT	CATCGCGGTC	CGGTCTCTCT	TAGGAGCGCG	AGTGTGCTAT	GTGGAAGCGC	1680
TGCGGAAGCG	GTGCGCGCGG	GGTGTGCGAA	TGATCTGCTGA	ATTGCTGCGC	GGATCTGCGCA	1740
CGGTGCTGCT	CGGTGCTGCG	GGGCGATGCA	GGTTGCGCGC	GTTCATGCTCT	CATCAGATGCG	1800
CTCGGCTGAT	CGGTGCAAGC	GGTTCGCGATG	TGCGGCTGCT	GAGCTACTTG	CGCGGCGAGC	1860
CGCGCAAGCG	GGAGGCGATC	TGCTGTGCGC	GTCTGCTGCT	GGCGGTGATG	GTGCTTGGCA	1920

TTATCGGCAC CACCACTACT GACCTGCTTC GGCAGCTGCC GGTGTTCGCC CCGGAGGGCG 1989  
CGATCGGGAA TTC 1993

## (2) INFORMATION FOR SEQ ID NO:155:

- (1) SEQUENCE CHARACTERISTICS:
  - (a) LENGTH: 374 amino acids
  - (b) TYPE: amino acid
  - (c) STRANDEDNESS:
  - (d) TOPOLOGY: linear

## (a1) SEQUENCE DESCRIPTION: SEQ ID NO:155:

Met	Lys	Ile	Arg	Leu	His	Thr	Leu	Leu	Ala	Val	Leu	Thr	Ala	Ala	Pro	
1			5						10					15		
Leu	Leu	Leu	Ala	Ala	Ala	Gly	Cys	Gly	Ser	Lys	Pro	Pro	Ser	Gly	Ser	
			20						25					30		
Pro	Glu	Thr	Gly	Ala	Gly	Ala	Gly	Thr	Val	His	Thr	Thr	Pro	Ala	Ser	
			35					40					45			
Ser	Pro	Val	Thr	Leu	Ala	Glu	Thr	Gly	Ser	Thr	Leu	Leu	Tyr	Pro	Leu	
			50			55					60					
Phe	Asn	Leu	Trp	Gly	Pro	Ala	Phe	His	Glu	Arg	Tyr	Pro	Asn	Val	Thr	
			65			70				75				80		
Ile	Thr	Ala	Gln	Gly	Thr	Gly	Ser	Gly	Ala	Gly	Ile	Ala	Gln	Ala	Ala	
			85						90					95		
Ala	Gly	Thr	Val	Asn	Ile	Gly	Ala	Ser	Asp	Ala	Tyr	Leu	Ser	Glu	Gly	
			100					105						110		
Asp	Met	Ala	Ala	His	Lys	Gly	Leu	Met	Asn	Ile	Ala	Leu	Ala	Ile	Ser	
			115					120						125		
Ala	Gln	Gln	Val	Asn	Tyr	Asn	Leu	Pro	Gly	Val	Ser	Glu	His	Leu	Lys	
			130					135				140				
Leu	Asn	Gly	Lys	Val	Leu	Ala	Ala	Met	Tyr	Gln	Gly	Thr	Ala	Lys	Thr	
			145			150				155				160		
Trp	Asp	Asp	Pro	Glu	Ile	Ala	Ala	Leu	Asn	Pro	Gly	Val	Asn	Leu	Pro	
			165					170					175			
Gly	Thr	Ala	Val	Val	Pro	Leu	His	Arg	Ser	Asp	Gly	Ser	Gly	Asp	Thr	
			180					185					190			
Phe	Leu	Phe	Thr	Gln	Tyr	Leu	Ser	Lys	Gln	Asp	Pro	Glu	Gly	Trp	Gly	
			195					200					205			
Lys	Ser	Pro	Gly	Phe	Gly	Thr	Thr	Val	Asp	Phe	Pro	Ala	Val	Pro	Gly	
			210			215					220					
Ala	Leu	Gly	Gln	Asn	Gly	Asn	Gly	Gly	Met	Val	Thr	Gly	Cys	Ala	Glu	
			225			230				235					240	



Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala  
245 250 255

Ser Gln Arg Gly Leu Gly Gln Ala Gln Leu Gly Asn Ser Ser Gly Asn  
260 265 270

Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Gln Gly Phe  
275 280 285

Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro  
290 295 300

Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Gln Tyr Ala Ile Val Asn  
305 310 315 320

Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu  
325 330 335

Ile Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val  
340 345 350

Ala Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu  
355 360 365

Ile Ala Thr Ile Ser Ser  
370

## (2) INFORMATION FOR SEQ ID NO:156:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1777 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

GGCTCTGACC ACCACCTGGG TGTCCAGCTG GGTGCGCGCA TTTAGTCCG GGTACTGCTG 60

GTTGGGGCGG GCGAACAAAT AGGACACAGC ATGCGACGAG CGTGGGTAGC GTTTACCGGT 120

GTAGCGAAAC GCGACGCGGG CCGCGTTTGG CACTTCTCTC AGCGCGTCTT TCGACACAC 180

CTCGTGGAAG GTGATGCGGT CGAATTGTG CCGCGGACAC CTGCGGACCA GCGCGATCG 240

CTGGACGCG GCGCGGCGCG TGTGTAAGCG GCATCCCGTT CAGCGGACG GCTTGCGCGG 300

CGCAAGCGAT ACCATTATTC GAACACCGGT TGTATATCTT GTCAACGCTG GCGCGTAGCG 360

AGTCCGCGAC AGGATGTGAT ATGACATCTC TGGCGGACCA GACAGGAGCG ATGCGTTATG 420

ACGACATTGG GCGTCGAGCT CTACGGGCGG CGCAAGTTCG TAGAAATCG GCGGAGCGCG 480

ATGCGGTATA TCGGCGAGGG CAGCGGTGAC GCGATCGTGT TTGAGCGCGG CAGCGCGACG 540

CGCTCTTACT TGTGCGCGCA CATCATGCGG CACTTGGGAG GGTGCGCGCG GCTGTGTCGG 600

TGCGATCTGA TCGGATGCG CGCGTGGAC AAGCTCAGCC CATCGGACC CAGCGCTAT	660
AGCTATGGCG AGCAACGAGA CTTTGTGTC GCGCTGCGG ACGGCTGGA GTCGCGGAC	720
CAGCTGGTAC TGATGCTGCA GACTGGGCG TGGGCGTCC GCTTGGATG CGCTAACGAG	780
CTTGGGAGCC GACTCGAGG GATCGGCTTC ATGGAAGCGA TCGTCAACCC GATGAGCTGG	840
GCGAGTGGC GCGCGGCGT GCGGGGTGTC TTCCAGGTT TCGGATCGC TCAAGGCGAG	900
CGATGCGCT TCGAGCACA CATCTTTTC GACCGGTGC TCGCGGCGC GATCTCGGA	960
CAGCTCAGCG ACGAGGAAT GAACACTAT CGCGGCGAT TCGGAGCGG CGCGGAGGAC	1020
CGTGGCGGA GTTGTGCTG GCGAGGAGC CTTCGATCG ACGGTGAGC CGCGGAGGTC	1080
GTGCGTGG TCAACGATA CCGAGGCTGG CTCGAGGAAA CGGAGATGG GAAACTGTT	1140
ATCAACGCG AGCGGCGCG GATCTCAGC GCGCGGATCC CTCGCTATCT CAGGAGCTGG	1200
CGGACGAGA CGAAGATCAC AGTCCGCGG GCGGATTCG TCGAGGAGGA CAGCGATGG	1260
GTGATATCT GCGCGGCGG TCGGAGGAT CGCGGAGCTG GAGCGGCTCT GATTTCAGA	1320
GACCAAGAT GTATTTGCG CGAAGGCGG CGCGGCTCT GTCAAGTCA AAGACTTCT	1380
GCTCGGCGA GAGATTCTC CGGAAAGGG CACCAATCG AGCGCTTTC TTGCAAGGA	1440
GCTGAGAAA TATAGTGGC AGGACAGGG TCTTCTATT TCGGAGGGA ATTACTGCT	1500
GCTTCTAT GCGTCACTT CAGGAGGCG GAGCGGATCA GCGTATCGG ATTGAGCTA	1560
TGGAAGGCT ATGATGAAG GTTGAAGCA TTGGAAGGAC GCGGCTCTCT GCGGCTGCG	1620
ATGATGCGC AGCGTCTCT TCGAAGGCG CAGATGACT AGCTGTACT AGCGGCGCG	1680
GCGATCGGG AGGAGTCTC GCGGCTCTG AGGAGGAGG GTATAGGCG GGTCTTCTAT	1740
TAGTGGCG TTGCGATTG GCGGCGGCG GTCGCT	1777

(2) INFORMATION FOR SEQ ID NO:157:

- (1) SEQUENCE CHARACTERISTICS:  
 (a) LENGTH: 124 base pairs  
 (b) TYPE: nucleic acid  
 (c) STRANDEDNESS: single  
 (d) TOPOLOGY: linear

(3) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GAGATTGAT CTGAGCTGT TCGTATGCG CTCGTTGAG TGAATGCGA TATCAAGCA	60
GCGATGTTT TGGGTGCGA GTTTGGGCG ATGCGGAGC GTTCTAAC CAGGTTTG	120
ATCAATATT CCGGGGAGC GTTGGGGA GCGGCGAGG ATGCTGCGA GCGGCGGCG	180
GCGGCTGCG CAGCGGAGC CTGATGCTC AGCGGCTG CAGGAGCTA GCGAGCTTT	240

GGGGGXTGTC GTCCACAGTG GTACTCCGGT GACGACCGGG GGGGTGGCT GGGTGAAGAC	300
CTTGACCGAC GCGCGGATT (TAG)	324

(2) INFORMATION FOR SEQ ID NO:158:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1338 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(2) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GGGTACCGCG CGGTTGAGT TGGCGGGGA CTTGTAGAC CTGAACCACT TGGCTGGCG	60
AAGGATTGAC GAACGCTG TGGGCGGCT GTGATGCTC AAGGTGTGG GTGATGCTG	120
CGATGACCGG CCGGGGACCC GGCCTAGCG CTTGGAAGAC GTCTTGCCG CCGCGACCGA	180
GCACGACTTC CAGCCCGACT CGATCGGCT GCTGACCGT CTTCTGCGT TGGTTCGCG	240
CGAAGCTGCG GTTGGGAGC GATTTCGTT CTTCACTGAC CTGAGCGCG ACCGAGCGC	300
GTGGCGCGCG TGGGAGAAC GCGACGCGG CGAAGTGGAG AACGCGCTG CGCTGCTGG	360
GTCTGTGTC ACCTTCGCG GATCTGGCG TTGCTGCGC ACGCTTGGG CTGACGCGG	420
CTGAATCGAC TAGATGAGG GATTGGGCG CGAATCGCG TGTGTGCTG ACCAAGACAC	480
GAGTACTGTC ATCTCTATT GATGCACTG ATGACCGCG TGAATLACCA GAGCAATCG	540
AAGTGGCGCG GCGAAACCT CTCGGAGATG TCGGCGTCC CTCGCGAAC CTGCGTGGT	600
GGGTCTATTC GGACATGCT CCGCTGCGG GATGCTGCT GACGCGCGG CTGAAGGAGT	660
CGAGCGCGCG GTTGCACCG CTGCTGAGC GCGGCGACG GTTACTGCTG GTGAAGCGCG	720
CGATGGCGGA GAAGCGTTT GAGGTGGCG GATCGGATT CTCTTTGTC CCGACGCTG	780
CGACACGCGA CCGGAGCGG GTTGGCGCG AGCAGCGCG CTTGCTGCG GCGGCGCGG	840
CTGACGCGAC CGAGGATGT GTCTACTCG GCGCGCGAC GAAGTTGTT CCGGCACTG	900
CGTTTACCG GCGGAGGCT CTGAGCGCA TCGAGGATCT GCACATCTG ACCGCGAGT	960
CGTGGCGCG GAGCGGCTC GACTTTCGG CTAAACACAA ACTGCGCGT TGGTGTGCT	1020
CGGGGATGCG GCTGCGGAG CGGTTCGCG TGGCGGTAG GCGCGGCTG GCGGCTTGA	1080
CGAGCTGGT GCGGCTGCG GTGAGCGCA CGTTCGCGG GCGGCTGCG GAGGAGCGG	1140
CGTGGCTGA GTTGGCTGCG GAGTTCGCG AGGCGGCTG TTGCTGCGG GGCATCGCT	1200
GAGTGTGAG TGTACGCGCA GTGCGCGCG CGAGTGAAT GCTGTGCTT CGGTGCTG	1260
TGGGTGAAT TGAAGCGCG GCGAACGCA GATTCGCGG CGCATCTG CCGCGCGCG	1320

GCGGCGACCG GTACACCG

1338

## (2) INFORMATION FOR SEQ ID NO:159:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 321 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:159:

CGCGGGGAC	CGCGGGGAC	GGCGGTACCG	GGCGGTACCG	CGGTGACGG	GCTGCTGTGG	60
TGGGCTTCGG	CGCGAAGGCG	GAGCGTGGGT	TGGGTGGGCG	CAAGGCGGT	AAGCGCGGAA	120
TAGGTGGGCG	CGCGGTGACA	GGCGGGTTCG	CGCGGGGACG	GCGACCGCG	GCGAAGGTTC	180
GCGCGGCGCG	TGCGGGGCG	GCGGGGACG	AGCGGGGAC	CACCGGAAAT	CGCGGGGTA	240
AGCGGGGCGA	CGCGGGGATC	GGCGGTGCG	CGCGGGGCG	CGCGGGGCG	GCGACCGGAA	300
AGCGGGGCGA	TGCGGGGAC	G				321

## (2) INFORMATION FOR SEQ ID NO:160:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 452 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:160:

GAGAGCGCG	CGCGGGGATA	TGCGGTGGCT	CGCGGTGGCT	TTTGGCGGAA	CGTGGCGCG	60
GGCGGTTCGG	GCTGATGATC	AGCGGTGGCT	AGCGGTGGCT	CGCGGGGAA	TGGGTGCTTC	120
CGTGGCTTC	CGCGGTTC	AGTGGGCG	AGCGGTGGCT	CGCGGGGAA	ATCGGTGGCT	180
CGCGGTTC	CGCGGTTC	AGCGGTTC	AGCGGTTC	CGCGGGGAA	CGTGGGTTC	240
AGCGGTTC	CGCGGTTC	AGCGGTTC	AGCGGTTC	CGCGGGGAA	CGTGGGTTC	300
AGCGGTTC	CGCGGTTC	AGCGGTTC	AGCGGTTC	CGCGGGGAA	CGTGGGTTC	360
AGCGGTTC	CGCGGTTC	AGCGGTTC	AGCGGTTC	CGCGGGGAA	CGTGGGTTC	420
AGCGGTTC	CGCGGTTC	AGCGGTTC	AGCGGTTC	CGCGGGGAA	CGTGGGTTC	480
AGCGGTTC	CGCGGTTC	AGCGGTTC	AGCGGTTC	CGCGGGGAA	CGTGGGTTC	540

## (2) INFORMATION FOR SEQ ID NO:161:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 276 amino acids  
 (R) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

```

Phe Ala Glu His Leu Val Gln Gly Asp Ala Val Gln Leu Trp Arg Ala
 1             5             10             15
Asn Ala Ala Asp Gln Ala Asp Pro Leu Gln Pro Gly Ser Ala Arg Arg
 20             25             30
Gln Arg Ala Ser Arg Ser Pro Arg Arg Leu Ala Gly Pro Asn Ala Tyr
 35             40             45
His Tyr Ser Asn Asn Arg Ser Ile Leu Cys Gln Arg Trp Pro Leu Pro
 50             55             60
Ser Ala Ala Glu Asp Val Ile Cys His Leu Cys Pro His Arg Gln Glu
 65             70             75             80
Pro Gly Leu Met Thr Ala Phe Gly Val Gln Pro Tyr Gly Gln Pro Lys
 85             90             95
Tyr Leu Glu Ile Ala Gly Lys Arg Met Ala Tyr Ile Asp Gln Gly Lys
100            105            110
Gly Asp Ala Ile Val Phe Gln His Gly Asn Pro Thr Ser Ser Tyr Leu
115            120            125
Trp Arg Asn Ile Met Pro His Leu Glu Gly Leu Gly Arg Leu Val Ala
130            135            140
Cys Asp Leu Ile Gly Met Gly Ala Ser Asp Lys Leu Ser Pro Ser Gly
145            150            155            160
Pro Asp Arg Tyr Ser Tyr Gly Gln Gln Arg Asp Phe Leu Phe Ala Leu
165            170            175
Trp Asp Ala Leu Asp Leu Gly Asp His Val Val Leu Val Leu His Asp
180            185            190
Trp Gly Ser Ala Leu Gly Phe Asp Trp Ala Asn Gln His Arg Asp Arg
195            200            205
Val Gln Gly Ile Ala Phe Met Glu Ala Ile Val Thr Pro Met Thr Trp
210            215            220
Ala Asp Trp Pro Pro Ala Val Arg Gly Val Phe Gln Gly Phe Arg Ser
225            230            235            240
Pro Gln Gly Glu Pro Met Ala Leu Glu His Asn Ile Phe Val Glu Arg
245            250            255
Val Leu Pro Gly Ala Ile Leu Arg Gln Leu Ser Asp Glu Glu Met Asn
260            265            270

```

```

His Tyr Arg Arg Pro Phe Val Asn Gly Gly Glu Asp Arg Arg Pro Thr
275                280                285

Leu Ser Trp Pro Arg Asn Leu Pro Ile Asp Gly Glu Pro Ala Glu Val
290                295                300

Val Ala Leu Val Asn Glu Tyr Arg Ser Trp Leu Glu Glu Thr Asp Met
305                310                315                320

Pro Lys Leu Phe Ile Asn Ala Glu Pro Gly Ala Ile Ile Thr Gly Arg
325                330                335

Ile Arg Asp Tyr Val Arg Ser Trp Pro Asn Gln Thr Glu Ile Thr Val
340                345                350

Pro Gly Val His Phe Val Gln Glu Asp Ser Asp Gly Val Val Ser Trp
355                360                365

Ala Gly Ala Arg Gln His Arg Arg Pro Gly Ser Ala Leu Ile Ser Arg
370                375                380

Asp Gln Glu Cys Asp Phe Arg Arg Arg Arg Arg Pro Ala Cys Gln Leu
385                390                395                400

Ile Arg Leu Pro Ala Pro Gly Arg Asp Ser Gln Gly Lys Gly His Gln
405                410                415

Ser Gln Pro Leu Pro Ser Gln Arg Gly Arg Gln Ile Tyr Val Ala Gly
420                425                430

Gln Arg Ser Ser Tyr Leu Pro Ser Gln Leu Val Ala Ala Phe Leu Trp
435                440                445

Ala Gln Phe Glu Glu Ala Glu Arg Ile Thr Arg Ile Arg Leu Asp Leu
450                455                460

Trp Asn Arg Tyr His Glu Ser Phe Gln Ser Leu Glu Gln Arg Gly Leu
465                470                475                480

Leu Arg Arg Pro Ile Ile Pro Gln Gly Cys Ser His Asn Ala His Met
485                490                495

Tyr Tyr Val Leu Leu Ala Pro Ser Ala Asp Arg Gln Glu Val Leu Ala
500                505                510

Arg Leu Thr Ser Glu Gly Ile Gly Ala Val Phe His Tyr Val Pro Leu
515                520                525

His Asp Ser Pro Ala Gly Arg Arg
530                535

```

## (2) INFORMATION FOR SEQ ID NO:162:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 284 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

```

Asn Glu Ser Ala Pro Arg Ser Pro Met Leu Pro Ser Ala Arg Pro Arg
 1           5           10           15

Tyr Asp Ala Ile Ala Val Leu Leu Asn Glu Met His Ala Gly His Cys
 20           25           30

Asp Phe Gly Leu Val Gly Pro Ala Pro Asp Ile Val Thr Asp Ala Ala
 35           40           45

Gly Asp Asp Arg Ala Gly Leu Gly Val Asp Glu Gln Phe Arg His Val
 50           55           60

Gly Phe Leu Glu Pro Ala Pro Val Leu Val Asp Gln Arg Asp Asp Leu
 65           70           75           80

Gly Gly Leu Thr Val Asp Trp Lys Val Ser Trp Pro Arg Gln Arg Gly
 85           90           95

Ala Thr Val Leu Ala Ala Val His Glu Trp Pro Pro Ile Val Val His
100           105           110

Phe Leu Val Ala Glu Leu Ser Gln Asp Asn Pro Gly Gln His Pro Phe
115           120           125

Asp Lys Asp Val Val Leu Gln Arg His Trp Leu Ala Leu Arg Arg Ser
130           135           140

Glu Thr Leu Glu His Thr Pro His Gly Arg Arg Pro Val Arg Pro Arg
145           150           155           160

His Arg Gly Asp Asp Arg Phe His Glu Arg Asp Phe Leu His Ser Val
165           170           175

Ala Met Leu Val Ser Pro Val Glu Ala Glu Arg Arg Ala Pro Val Val
180           185           190

Gln His Gln Tyr His Val Val Ala Glu Val Glu Arg Ile Pro Gln Arg
195           200           205

Glu Gln Lys Val Ser Leu Leu Ala Ile Ala Ile Ala Val Gly Ser Arg
210           215           220

Trp Ala Glu Leu Val Arg Arg Ala His Pro Asn Gln Ile Ala Gly His
225           230           235           240

Gln Pro Ala Gln Pro Phe Gln Val Arg His Asp Val Ala Pro Gln Val
245           250           255

Arg Arg Arg Gly Val Ala Val Leu Lys Asp Asn Gly Val Thr Leu Ala
260           265           270

Phe Val Asp Ile Arg His Ala Leu Pro Gly Asp Phe
275           280

```

## (xii) INFORMATION FOR SEQ ID NO:163:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 264 base pairs

167

(W) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:163:

ATGAACATGT	CGTCGGTGT	GGCTGCGAG	GGCTTTGGGC	GATTCCGCG	CTAATCTCTC	60
GGCATGACAG	CGATCCGCG	TTTCTCCGAT	GGCTTCGCG	AAGAGTGGC	GGGTAGCGGA	120
ATCGCGGTCT	CGGTGATCCA	CCTGCGCTG	ACCGAGACG	CCTGTTTGG	CAAGTGGAC	180
CGCGCGGACA	TGCGCGCGC	GTTCGCGAC	CTCAGCGCA	TTCCGCTTCA	CTGGGTGGC	240
CGAGCGGTGC	TTGACGTTGT	GGCG				264

(Z) INFORMATION FOR SEQ ID NO:164:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1171 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:164:

TAGTGGCGGA	GAATGACCTC	CGGTTCCAG	CGACCCGCTT	CGAGTACAG	CGGAGCCAG	60
AAGCCCGTGC	GATCCTTACC	CGGAAAGCAG	TGGTGGCA	CGGAGCTCC	CGCGGAGGC	120
AATGTAGCA	CAGCATGTAG	CGCGGCTGT	CTCTCATTC	CGCTTGGAA	TTGGCGATAC	180
TGCTGCTCA	TGTACCGGT	GGCGGCTCA	TTATCGACT	GGCTGGTTG	GGGAGCTTC	240
CGCTTGGACT	CTCATTTGCT	TAGCAAGCTC	TTGATTCGG	TTCTGTGGC	CGCTGAGTC	300
TGCGGCTCAT	CATCGCGAG	GTGGGGGAC	GGCAAGAGT	GGAGTGCAT	GGCTTCGGA	360
AGCGGCTCTG	GACCGCGG	GGCAAGCTC	CGAGAGAGC	CGAGTGGGC	AAGCTCGGTG	420
ATCCGAGCC	GGCGAGAGT	TGCGGCTGT	GGGAGTTG	CGAGAGAGT	GGCGAGAGC	480
CGCGGCTC	CAAGCAGCG	TTGGGAGTA	CGGATCTCA	CTTCGCGATC	CGCGAGCCA	540
ATCTCTCTGC	CGGCACTCT	CAGATCCCG	TGCTGGTTC	ACAGAGAGG	CGCGAGATCT	600
CGCGAGGCT	ATCGGAGAT	GACCGGCA	CGCATTTCT	CAATCGCTGC	GGCTTCGCG	660
ACTATTGCA	CTTCCGCG	GTGGGCTAT	TAGGAGCA	TGGAGCTTC	GACGAGCTG	720
CGCGAGTAA	CGCGAGGCT	AGCTCCGCG	GTGAGCGGA	GGATCGCGG	GTGATCTTTG	780
CCTGCGAGT	CGTACGCTT	GATCGAGCTC	TTGCGGTCG	GGCGGCGAG	CGCGATCGC	840



168

TTATCGACCT CGGGTATGCG CAGCGGCAAG CTGGGGGCGT TCGTCGAGGT CAGGAACTCC	908
ACCATCGGCA CGGCGACGCA GGTGCGCGAC CTGACCTAGG TGGCGGACGC CGACATCGGC	966
GATACAGCA ACATCGGCGG CTGACGCGTA TTTCTCAACT ACGACGGTAC GTCTAAACGG	1020
CGGACGACCG TCGGTTCGCA GTTACGCGAC GGTTCGACA CATTCTTCCT GGGCGGACGA	1080
ACCATCGGCG ACGGCGCGTA TACCGGGGCC GCGCGGTGG TCGCGGAGGA TGTCCCCCGG	1140
GGGGGCTCGG CATTGTGCGC GGTTCGCAA C	1171

(2) INFORMATION FOR SEQ ID NO:165:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 227 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) SEQUENCE DESCRIPTION: SEQ ID NO:165:

CGAAGGGCGG CAGGCGCGG CCGCGGCTGA ACAGCGTGGG CCGGTGCTA GCGGCGCAAG	60
ACGCGCGCTA AGCGGCGAC GCGGCGACCG GCGCGACGC GCGGCGCGG GCGACGAGCT	120
TGACCTAAGG GCGGCGCGG AACCGCGCA ACGCGCGTGA GCGCGGCTC GCGGCGACG	180
GCGGAACCG GCGGAACGC GCGGCGACA CCGCGCGCG CCGCGCG	227

(2) INFORMATION FOR SEQ ID NO:166:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 304 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) SEQUENCE DESCRIPTION: SEQ ID NO:166:

CTGTGTGTC ATGCGCGCG AGGGGCTAG CGTGGCGG GGTCTACCC CAGGCGCAA	60
GGGCGCGCG GGTCTGCTC CAGCGCGCG GCGCGACCG GCGCGCGCG GCGAGGCGG	120
CAGCTCCGA GTGTGCGCG GCGAGCGCG GCGCGCGCG ATGTGCGCA AGCGCGCGG	180
GCGCGCGCG GCGCGCGCG GCGCGCGCG GCGCGCGCG GGTTCGCG GGTGAGTGC	240
CAGCGCGCG AACCTGCTG AAACGCGCG AACGCTAAC CCGCGCGCG AGGTGCGCG	300
CGCG	304

(2) INFORMATION FOR SEQ ID NO:167:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1439 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

GTGGGAGGCT GCGGAGGCTG TATAACAAGG ACAACATGGA CAGCGGCGG CTGGTGAAG	60
TGATCAGGCT ATTAAACAGT GCGCGCTTCA GCGGCGAGGG CAGGACATCG GCGCGGAGTC	120
TGATGAGTGA GCTCTACGAA TACTTCTCTG GCAATTTCGG TCGCGAGGAA GGGGAGCGGG	180
GTGGCGAGTT CTTTACCTCG CCGACGCTGG TCAAGGTGAT CTTGAGGTGG CTGAGGCGGT	240
GAGTGGCGCG GGTGTATGAC CGTGTCTGG GTTCCGAGG CATGTTTGTG CAGACCGAGA	300
AGTTCTATCTA CGAACAGGAC GCGCATCGGA AGGATGTCTC GATCTATGGC CAGGAAGGCA	360
TTGAGGAGAC CTGGCGGATG GCGAGATGTA ACGTCGGCAT CGACGGCATC GACAACAGGG	420
GGCTCGCGCG CCGATCGAGT GATACCTTGG CCGCGGACCA GCACCTGGAC GTGCAGATCG	480
ACTACGTGAT GCGCAATCGG CGGTTGAGCA TCAAGAGCTG GCGCGGCAAC GAGGAAGACC	540
CAGCGTGAGG CTTCGGTATT GCGCGCGGCA ATAACTGGAG CTAGCGATGG ATTGAGCACA	600
TGCTGTACAA CTTCGCGGCG GAGGCTCGGG CCGGCGTGGT GATGCGCAAC GGGTCGATGT	660
CTTGGAATTC CAACCGCAGG GGGGATATTC GCGCGCAAAI CTGTGAGGCG GATTTCGTTT	720
CTTGATAGGT CGGTTTCTCC ACCGAGTGTG TCGCGACGAC CGGAAATCCG GTGTGCTGT	780
GCTTTTTCCT CAATAACAAG GCGGAGGSTR AGCAAGGCTC TATCAAGTGG TCGCGGCGAG	840
TGCTGTTCAT CGAGCTCTGT GAACTGGGCG AGCTAGTGGG CCGGAGCGAG CCGGCGCTCA	900
CGAACGAGAA GATCTGCGCC ATCGAGGATA GCTTCGATGC GAGCAGGAGC ACCGCGAAGC	960
CGGCTCGCGG TGCTGAGGCG GGTATGCGGG GCACGTGAGT CTACCGGCGG GCGGCTGCTG	1020
GCGGCGCGCG CGGCGAGCGG GCTGTGCGCG GCTGTGCTTT GAGCAAGGCT GTGGCGGCGG	1080
ACGCGCGGCA GCGCGGCAAC GCGGCGGAGG GCGGCGGAGG CACCGAGGCG GCGGCGGCGG	1140
GCAAGCGCGG CAAGGCGAGG AGCGGTGCGG CCGCGGCTG AGGCTGCTG AAGTTCACCG	1200
CGGCGGAGCG GCGCAAGGCG GCGAATGCGG GCAACGCGCG CACCGCTCCG GCGGCGGCGG	1260
GCGCGGAGCG GGTGTGCGCG GCGAGGCGCG GCACGCGCGG CCGCGGCGCG GGTGCGAGCG	1320
GCGCGGCGAG GCGCAAGGCG GCGAAGGCGG CCGAGGCTGC GCGGAGGCG TCAAGGCTCA	1380
TCAAGGCTCA GCGCGGCGAG GCGGCGAGCG GCGGCGGAGG CCGCAAGGCG GCGGCGGCG	1439

(2) INFORMATION FOR SEQ ID NO:168:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 329 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

GGGCGGGGGG GGGGCGATT TCTCGTGCT TGAATGTCGG TGGGGATAAC GGGCGTGATS	60
GGGATACGG CCGGATGGG GGGCTGGGG GGGCTGGGG CCGCGGGGG GCGCGCGGG	120
TGATCGGCT GCTGGGGGG CAGGGGGGG GCGGGGGGG GGGGACGGG GGGGCGGGG	180
GTGTGGGG TGGCGGGGG GCGGGGGGG GCGGGGGGG GCGGTTCAC GCGGTGGGG	240
GCGGGGGGG GCGGTTCAC GCGGTTCAC GCGGTTCAC GCGGTTCAC GCGGTTCAC	300
CGGGGGGG GCGGTTCAC GCGGTTCAC GCGGTTCAC GCGGTTCAC GCGGTTCAC	329

(2) INFORMATION FOR SEQ ID NO:169:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 80 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

CGACCGTGG CAGGGGGGG AGGGGACCA GCGTGGGGT GCGGGGGT ATCTGCGG	60
CGGGGGGGT GATCGGCAAC	80

(2) INFORMATION FOR SEQ ID NO:170:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 392 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

GGCTCTGTC GACTACAC CCGGCGATT GGGGACGTC GGGGCGATT ATCCAGCTCA	60
AGGCTACTA GTTACGTC CAGGACGGG GCGTACGCT GGGGTCAGC GCGGAGGAA	120
TGAGGTCAT CAGGACGGG GGGGACGTC GGGGACGTC GGGGACGTC GGGGACGTC	180

171

```

CCCCGCGCA CTTCCGGGCG CAAGCGGCTT CAGCGCTGCG AACGGGGGCG ATCTGTGTAG 240
CACACTGAT GGGGGGCAAC GAGATTGCTC CAATTGTCAA GCGGTGTTCG ACCGGAGGGA 300
CGGTTATAC GATATTCAGC CTATGTGACT GCGAGAACG GGCATAACGA TCCCGTGATC 360
CGCGGACAGC CCACGAGTGC AAGACGGTGA CA

```

(2) INFORMATION FOR SEQ ID NO:171:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 535 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

```

ACCGGCGCA CCGGGGCGAC CGGCTTCGCC GGTGCGGCG GGGGGGCGCG CCGGACGGGG 60
GGTATCAGCG CTCCGGGCGG CACCAACGCG TCTGCTGGCG CTGGGGGAC CCGGGGACAA 120
GGCGGGGCGG CGGGGCGTGG CCGGGGCGCG GCGGATAAC CCAATGGCAT CCGGGGCGCC 180
CGCGGCGACG CCGGCGCGCG CGGAGCGGCG GGAACCGGCG GGGCGGTCGG CCGGATCGGT 240
ACCGGCGCA CCGGGGCGAC GGTGGGCGCG CTGGGTACCG CCGGATCGCG CGGTACCGGG 300
CGTACGGGCG GTCTCGGTGG TGCTGTGTGT CGAGGTGCGG CTGGGGGCG TGGCGGACCG 360
CGTACCGGCG CGCGCGGCTT CAGCGCGCGG GCGGGCGAG AAGCGGACCG GGGCGGCAAC 420
AGCGGTGTGG CGCGGCACTA CCGCTCGGCG GGGCGCGGCG CTGGAGCGCG CAGGCGCGCG 480
ACCGGAGGTG CCGCGCGGTC CCGCGGGGAC AACCGGACCG GTGTTGCTTT CGGCG 535

```

(2) INFORMATION FOR SEQ ID NO:172:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 699 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

```

CGGAGTGGCG CCGGGCGATA CCGGGGTCA CAGCTACTAC ATCATGCGCA CCGGGAATCG 60
GCGCGTGTGG CAAGCGCTGC GGGCGGTGCG GGTGATCGCG GATCGCGTGG CCGAGCTGAT 120
CGAGCGGACG CTGAATGTGA TGCTCAAGCT GGGCTAGGCG GACCGGACTT AGGCTACTTC 180
GAGAGCTACG GCGATGTGCG GAGCGGCTTT CCGGCTGTGG CCGAGCTGCG CCGGCTAGGT 240

```

GATGGCGAT GCGCTGGCCS CGGAGACAC AGAAGGCATC CTTGACTTCA CGGCGGACCT	300
GCAGGCGCTG TCGGCGCAC CAGTCAGGCT CGCGCAGATC CAGCTGCTGC AACCGCGGCA	360
TCTGGTGGCC GGGGTGGGCG CCGACCGGAC GCGGCGGAG GTGGTGACA CAGTCGCGAC	420
GATCATCTCA ACCAACTAGC CGTCTCTGCT GCGGACGGTG GACATCGCCG TCGCTGGTCT	480
ACCAGCTCG CCGTGTACAC CAGCCCACTG TTGTCAGGC AACTCGGTGG GGGCAATCTG	540
ATCAAGGGA TCGGTATCG CCGGCGGCC ACGTAGGTT TAGGCAAGAT CGATAGCGG	600
CGGCGTGAA TTGCTACCC TCGTCGCGCG GCGTCGGAC AGGCTTCGAA ACAICGAGG	660
CGTCGTACG TAACGGATTC CCGACGGCAT	690

(2) INFORMATION FOR SEQ ID NO:173:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 407 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:

AGGTGACGG GAGTATCGG CGCGGCGAC GCGGACAGG CGGGAATCGG GAGTGGCTCT	60
TGGGACAGC CGGGGTGGC GCGAAGGCTG GCGGCGGAG CAGCGTACT CGAGGTGGG	120
GCTGTGGCG CACCGGCGG GACGCGGGA CCGGCGGCG TGGCGGCTG TTAATGGCG	180
CGGCGCGCG GCGGCGGCT AGCACTGGG GCGCGGCGG TCGCGGTGC GACGTTGGG	240
GCGGCGGCG GCGCGGCGG GCGCGGCGA CCGGCGGCG CGGCGGTCA GCGCGGCTG	300
TGTCGCGCG CGGCGGCGG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG	360
CGGCGGCTT CGGCGGCGG ATGCGCGGCG TGGGTGGTAC CGGTGAC	407

(2) INFORMATION FOR SEQ ID NO:174:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 466 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GATCGCTCG CGATCGGCG TCGCGGCGA GCGATTGCG GTCTCAGCG AAGACATCG	50
TGACGCGCG GCGCGGCGG AGCGGCTGC GGTGCGGCG GTGCAAGCG TCGGCGGCG	120

ACAGGCGAGTC CTTCGGGGCC TACGAGGCCA ACACATCCAT GTACCGGCTG TAGATGAGCG	180
GGATGCGGCG CTCCKCCACG GCATTCCGCG AGCGCCGCGC GTCTTTGTGA TGCTCGACGA	240
TCAGCGGAT GTCTGGGCG ACCAGCGGCG GCGCGGGAA GGTGCGGCG CTGCGCAGTA	300
GCGCGCGAC GTTCGGGCG AGTGCTCG CGATGTCCG GCGCAGCGT CGGCGCGAC	360
GCGCGAALA CGAGCGTCA CCGAGCTGG TCCGCTGG ATATCCCTTC CCGCTTGGC	420
CGATATTGA CGCGATGCG CGAGCGCGT ACAGCGCGC CACCAAGC	480

## (2) INFORMATION FOR SEQ ID NO:175:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 219 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:175:

GCTGTAAGC GCGGCGAGG TGGCATCGCG GCGCGGCG ACAGAGCGCG CGAGCGCGCG	60
GGGCGAATG CTACCGCG AAAGCGCG AGCGCGCTA GCGTGGTAA CGTGCGGAC	120
GCGCGCGCG CGCGATAG GCGCGGCG GCGAGCGCG AGCGCGCG GTACCGGAC	180
GCGCGAGCG GCGCGCGCG CGAGCGCG AAGCGCGC	219

## (2) INFORMATION FOR SEQ ID NO:176:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 434 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:176:

TAGCTCGCG GAGCGCGCA AGAGCGCGA CGTTCGCG GCGGTGCG GCGTTCGCG	60
CACAGTTCC GTACCCAG CGCGCGCG CGTTCGCG GCGCGCGCG GCGCGCGCG	120
CAGCGCTTT TTGCGCGCA AGCGCGCTT CGCGCGCG GCGGTGCG GCGGCGGAA	180
CGCGCGGCT AGTTCGCG CGTTCGCG TGCGCGCG AGCGCGGT TGCGCGCG	240
GGCGCGGAC GCGCTTTT CGGTTCGCG CGCGCGCG GCGTTCGCG GCGCGCGCG	300
CAATGCGCG GCGTTCGCG GCGCGAGCG CGCGTTGCG GCGCGCGCG GTGCGCGCG	360
CAGCGCGCG GCGTTCGCG AGCGCGCTT CAGCGCGCG AGCGCGCG AGTTCGCG	420

GATCGGATCA GGGCGCTACT GCGCGGAGCA GGTGGACAT CACCCGAGTT AGCGGGGGC 480  
ATTTCCTGAT CACC 494

## (2) INFORMATION FOR SEQ ID NO:177:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 220 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:177:

GGGCGGTTG TGGCGGGGC CAGCTCTTCA GCGCGGAGG GCGGGGGGT GCGTTTGGG 60  
TTGGGGGAC GCGGGGACG GGTGGGCTG GCGTGGGG AGGAGCGGG GCGGAGGCC 120  
CGCGAGCAC AGTGTGACC GGTGTATCG GTTGGCTGG GGGGGCGGC GCGGCGGGG 180  
GCGGAGCGG CAGCGGATT GCGGGCGCA TCAAGGGTC 220

## (2) INFORMATION FOR SEQ ID NO:178:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 358 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:178:

ATGGCGGCA GGGGGGCGC GCGGCTGTC GCGGGGCGG CAGTACAT TTCCAACGC 60  
GGGCAGGTT GTGCGGGCG CAGAGCGGG CAGAGCGGC TGGCGGGGC AAGACCAAC 120  
TGATCGGCT AGCTGCACG GGAAGGCGG ATGACACG CAGGATGCC GCTTCCTTG 180  
CGGCTTGGC CAGAGCGGG ATACCTTAC CTGACCGAG CAGGCGCTA AGCGCGGCA 240  
AGGAGATGC TGGCTGTAT GGTACCGGG TACAGTCT ACAGCTGTT GCGGAGCTC 300  
GGGACTACA TCGGAGGTC AGCATGACA GCGGGGCAA GTTCTCTGC ATGACATAC 360  
GCGCTATTC GCGGAGACG CTGAGCA 396

## (2) INFORMATION FOR SEQ ID NO:179:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 400 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (11) SEQUENCE DESCRIPTION: SEQ ID NO:179:

GCAAGAGCGG	CAAGGCGCGG	GCGGCGATGA	ACAGGCTCGA	CGCGCTGCTA	GCGGCCCAAG	60
ACGGCGGCGA	AGCGGCGCGC	GCGGCGACCG	GCGGCAACCG	CGCGCGCGCG	GCGACGAGCT	120
TGACCTCAAG	CGCGGCGCGC	AACGCGCGCA	ACGGCGCTGA	CGCGCGGCTC	GCGGCGAAGC	180
GCGGAACCGG	GCGAAGCGCG	GCGGACACAA	CGGCGACCGC	CGCGCGGCGC	ACGACAGCGG	240
GCGAGCGCGG	GCGGCGGCGG	GCGGCGGAAA	CGGCGGGAAC	CGCGGAGCGC	GCGGCGACCG	300
GCGCGGCGCG	CGAACAAGCG	ACGGCGGCGA	ACGGCGGCGA	CGCGCGCAAA	GCGGCGACCG	360
GCGGCGACCG	TGACCTCTCA	GCGGCGACCG	GCGGCGCGCG			480

## (2) INFORMATION FOR SEQ ID NO:180:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 538 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (11) SEQUENCE DESCRIPTION: SEQ ID NO:180:

GCGACGCGCG	GCGACGCGCG	CATCGCGCGG	ATTGCGCGCG	ACCGCGGCTC	GCGGCGCGCG	60
ACCGCGCAAG	GCGGCGCAAG	GCGGCGCGCG	GCGGACCGCG	GCGAAGCGCG	GCTGCGCGCG	120
CAACAGCGCG	ACCGCGCGCG	GCGGCGCGCG	TGCGCGCGCG	ACCGCGCGCG	GCGGCGCGCG	180
GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	240
CACCGCGCGG	CACCGCGCGG	ACCGCGCGCG	ACCGCGCGCG	ACCGCGCGCG	ACCGCGCGCG	300
GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	360
GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	420
GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	480
GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	GCGGCGCGCG	538

## (2) INFORMATION FOR SEQ ID NO:181:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 239 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear



## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:181:

GGGCGGATGG TGGCGGGGAG GAGCTCTTCA GCGCGGAGG GCGGCGGAT GCGCTTGGGG	60
TTGGCGGCAC CGGCGGGGAG GGTGGGGGTG GGGGTGCGGG AGGGGGGGGG GCGGA/GGCG	120
CGGCGGAGAG AGGTCTAAGC GGTGGTACCG GCTTCGGTGG GGGGGCGGG GCGCTCGGGG	180
GCGCGGGGGG CAAGGCATTT GGGGGCGGCA TAAAGGGCTC CGGTGGTGGG GGGGGCACC	240

## (2) INFORMATION FOR SEQ ID NO:182:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 995 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:183:

AGCAGCGCTA CCGGTGGGGC CGGTTTGGCC GGGGGGGGCG GCGAGAGAGC GGGAGCGGGC	60
GGCAACAGCG GTGTGGGCGG CAGCAAGCGC TCGGAGCGCG CCGCGGGTGC AGGGGGCAAG	120
GGCGGACCGG GAGGTGGCGG CGGGTGGGC CGGACAGAGC CCATCGGTGC TGCTTTCGCG	180
GCTTGCCTCG GCGGACAGG TGGGGCGGGC GGGGGGGGCG GGGCGGGGCG GGGAGCGGT	240
ACCGCGGCGA CGGCGGGGT TGTGGGGGCG ACCGTAGTGC CAGGATCGGG CCGGGCGGGC	300
GGCGGGGGCG GTAGCGGGG CGATGGGGCG AGGGTCTCG GCTGGGGCT CTTCGATTT	360
GAGGGTGGCG AAGGCGGCGA AGCGGGGGCG GGGGGGAGCG CCGGCGGGCG GGGCATCAAG	420
GGGGCGGGCG GGGCGGGGCG CAGGGGGGCG GAGGGGGGCG ACCGGCGAAG CGGTGCGGCA	480
GCTCTGGGCG ACAGCGGGCG GGTGGGGGCG GAGGTGGGCG CGGTGGGGCG GGGGGGAGCG	540
GGGGGAGAGC CGGGGCTGCG GGTGACAGCG AGGGGTGGCG ACCGGGGGCG GGGGGGCAAT	600
GGGGGAGAGC GGGGGGCGCG CGGTGGGGCG GGGGGGCGCG ACAACATTT CAACGGGGCG	660
CAGGTGGGCG CGGGGGGCGA AGGGGGGCGA GGGGGGCGCG GGGGGGCGCG CAGGAGGTGG	720
TGGGGTGGCG CGGACCGGGG AAGCGGGGCG CAGGAGGCGA GAGTGGGGCG TTCTTGGGCG	780
GCTTGGAGCA GGGGGGCGCG AGGTGGGGCG AGGGGGGCGA GGGGATAGCG GGGGGGAGCG	840
GAGTGGGGCG GGTGGGGGCG AAGGGGGGCG CAGGATAGCG GGTGGGGGCG GAGGGGGGCG	900
AGTACAGTGC CGGGGTGAGC ATGGAGAGCG GGGGGAGGTT GGTGGGGGCG GATGAGGGCG	960
GTTAGTGGCG GAGACAGGCG GAACA	995

## (2) INFORMATION FOR SEQ ID NO:183:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2136 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:193:

CGGCAAGAGG	ATGGGTATCC	CCCGGCATCG	GGAGTGGGCG	ATTGCGCGGG	TTTGGCCAGC	60
GCAGCAAGAC	CGTACGAGA	TTGGGCTGCC	GAAGTAGGGC	GATCGGTCCG	CGATGCGGGC	120
ATGAAGGAGC	GCCATCAAT	TATGCGCGGA	ACCTTTCAGT	TTAGCGAGCA	TATGGGTAT	180
AGCACTAAGC	AGGATGATC	GATATGAGCC	AGTCCGAGAC	CCTCGGGTGC	GATCAGCAAG	240
AGATTTTGA	CAGGGCCAG	GAGTGGGAGG	CCCGGATGCG	GGACCGACCG	ACTGATGTCC	300
CCATCACGCG	GTGGGACTC	ACGGGGGCTA	AAACGCGCGC	CTAACGCGTG	GTAATGTCCG	360
CCGACAACT	GGCGAATAC	CTGGCGGCCU	GTGCCAAGAG	GGGCGAGCGT	CTGGCGAGCT	420
CGGTGCGCA	CGCGGCGAG	CCGTATGCGG	AGGTGATCGA	GGAGGTGCGG	ACCGGCGCTC	480
ACAAGCAGCG	CGAGGGAATC	GTGCGGCGAG	AATCGGCGCG	GGCGGTGCGA	GGGCGAGCTT	540
CGCGCGAAT	AACGGATAGC	CGGAGGCTCG	CCAGCGCGCG	TGAACCGAGC	TTCAATGCTC	600
TCAAGGAATC	GCATAGGAAG	CTCGAAGCGG	CGACCAAGAG	CGCTTCGCTC	GGCGACTTTG	660
CGGATGCGTG	GAAGACTTTC	AACGTGAGCG	TCAAGCGGGA	GGTCAAGCGG	TTCCCGGCGCT	720
TTGCGAAGTC	CGAGGCGGCT	GGGCTAGCGG	CTTGGAGGCG	TTGGCTCGGT	CAACAACCGC	780
AATGGATACT	CGCATGCCC	AATTTAGCGG	CTGGGATGCG	CAGCGAGCGT	CAATATCTTC	840
CGCGCTTCA	CGTTGGGCT	AGCGGGGAAC	ATCGGACTTA	TGAAGCAAT	GTGCGGCTCG	900
AGCGCTTCA	CGGGAAGAG	CTTTGCGGCG	GGGCGCAAT	TCTCCGCGTG	TACGCGGAGT	960
ATGCGAGAG	GTGCGAGAG	GTGCTGAGCG	AATACACAA	CAGGCGAGCG	CTGGAGCGCG	1020
TAAACCGGCG	GAAGCTTCCG	CCCGGCAATCA	AGATCGAGCG	CGCCCGGCGT	CGCGAAGAGC	1080
AGGATTTGAT	CGCTGCTTTC	CTGATGCGCG	CGTCTGAGCG	CTCGGCTGTC	ACTCGCGGTA	1140
CGGGATGCG	AGCGGCGCGG	ATGTTTCCCG	CTACCGGATC	CGCGGCTGCT	GGCTTCCCGG	1200
CTGAGCGCG	GGCGGAGCTG	AGTGGGCTG	GGGCGGAGCG	GCGAGGCGTG	TGGGCGGAGG	1260
TGGGCTGCGA	AGCGGATCTG	CTGCTGAGCG	CTGAGCGGCG	CGGAGTCCCG	TGGGCGGCGT	1320
TGGGATGCG	GATCGGCGCG	GGCGAATCTG	TGGGCGCGCG	TGGGCTGCTG	GACATTGCGG	1380
GCTTAGGCGA	GGAGAGCGCG	GGGCGGCGCG	CGCGGCTGCG	CGCGGCTGCG	ATGCGAATCTG	1440
CGATGGGCTG	CGCGCATGAG	GGACAGGCGG	GGCGGAGCTT	CAAGGCTTCT	CAGCGAGGAG	1500

```

ACGAGGAGCT CTACACCCAG GATCGGGCAT GGGCCGAGCC CGTCATTGGT AACCATCGGC 1560
GCCAGGACAG TAAGGAGTCC AAGTGAGCAT GGAGGANTTG GACCCGCATG TCGGCGGGCC 1625
GTTCAGACTG GTCGGCGGGT TTAGCTGGG CTTAGACGGG AGCTTGAATC AGATGAACAA 1680
CGGATCTTTC CGGCGCACCG ACBAGCCGA GAGCTGGA GTACAGATCA ATGGGCACTA 1745
GTGGCTCACC GGGCTGCGCA TCGAAGATGG TTGCTGAGG AAGCTGGGTC CGAGAGGGGT 1800
GGCTCAGGCG CTCACGAGG GCTGACACAA TGGCGAGCC GGGGCTGCG GGTATACGA 1860
CUGGGCGGCC GAGCAGCTGA CCGGTGCTT ATGGGCTAG TCGCGCGGCA TGAACCAAGG 1920
AATGGCTAA GCGCATTCCT GGGTGGTAG GAACTACGA GGAATGAGC GCGGCAATGC 1980
GGTCATTGAG CGGCGCGGAG ACGGCTGAG TACGATTGT CAGTGTTC ACATGGATCG 2040
CGCGGTTTGG GAGGGGCTTA TATGCTGGT CAGCAATATT GCGGAGCTA GCTGTCTTAA 2100
GGTGGGCTTA GCTGGCTTAA TTATGACTC CATTACCA 2136

```

(2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:

- (a) LENGTH: 468 amino acids
- (b) TYPE: amino acid
- (c) STRANDEDNESS:
- (d) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

```

Met Thr Glu Ser Gln Thr Val Thr Val Asp Gln Gln Glu Ile Leu Asn
1      5      10
Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro Thr Asp Val
20     25     30
Pro Ile Thr Pro Cys Glu Leu Thr Ala Ala Lys Asn Ala Ala Gln Gln
35     40     45
Leu Val Leu Ser Ala Asp Asn Met Arg Glu Tyr Leu Ala Ala Gly Ala
50     55     60
Lys Glu Arg Glu Arg Leu Ala Thr Ser Leu Arg Asn Ala Ala Lys Ala
65     70     75     80
Tyr Gly Glu Val Asp Glu Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly
85     90     95
Glu Gly Thr Val Gln Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser
100    105    110
Ser Ala Glu Leu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro
115    120    125
Asn Phe Met Asp Leu Lys Glu Ala Ala Arg Lys Leu Glu Thr Gly Asp
130    135    140

```

Gln Gly Ala Ser Leu Ala His Phe Ala Asp Gly Trp Asn Thr Phe Asn  
 145 150 155 160  
 Leu Thr Leu Gln Gly Asp Val Lys Arg Phe Arg Gly Phe Asp Asn Trp  
 165 170 175  
 Glu Gly Asp Ala Ala Thr Ala Cys Glu Ala Ser Leu Asp Gln Gln Arg  
 180 185 190  
 Gln Trp Ile Leu His Met Ala Lys Leu Ser Ala Ala Met Ala Lys Gln  
 195 200 205  
 Ala Gln Tyr Val Ala Gln Leu His Val Trp Ala Arg Arg Gln His Pro  
 210 215 220  
 Thr Tyr Glu Asp Ile Val Gly Leu Gln Arg Leu Tyr Ala Glu Asn Pro  
 225 230 235 240  
 Ser Ala Arg Asp Gln Ile Leu Pro Val Tyr Ala Gln Tyr Gln Glu Arg  
 245 250 255  
 Ser Glu Lys Val Leu Thr Glu Tyr Asn Asn Lys His Ala Leu Glu Pro  
 260 265 270  
 Val Asn Pro Pro Lys Pro Pro Pro Ala Ile Lys Ile Asp Pro Pro Pro  
 275 280 285  
 Pro Pro Gln Gln Gln Gly Leu Ile Pro Gly Phe Leu Met Pro Pro Ser  
 290 295 300  
 Asp Gly Ser Gly Val Thr Pro Gly Thr Gly Met Pro Ala Ala Pro Met  
 305 310 315 320  
 Val Pro Pro Thr Gly Ser Pro Gly Gly Gly Leu Pro Ala Asp Thr Ala  
 325 330 335  
 Ala Gln Leu Thr Ser Ala Gly Arg Glu Ala Ala Ala Leu Ser Gly Asp  
 340 345 350  
 Val Ala Val Lys Ala Ala Ser Leu Gly Gly Gly Gly Gly Gly Val  
 355 360 365  
 Pro Ser Ala Pro Leu Gly Ser Ala Ile Gly Gly Ala Glu Ser Val Arg  
 370 375 380  
 Pro Ala Gly Ala Gly Asp Ile Ala Gly Leu Gly Gln Gly Arg Ala Gly  
 385 390 395 400  
 Gly Gly Ala Ala Leu Gly Gly Gly Gly Met Gly Met Pro Met Gly Ala  
 405 410 415  
 Ala His Gln Gly Gln Gly Gly Ala Lys Ser Lys Gly Ser Gln Gln Gln  
 420 425 430  
 Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Glu Ala Val Ile  
 435 440 445  
 Gly Asn Arg Arg Arg Gln Asp Ser Lys Glu Ser Lys  
 450 455 460

180

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 277 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(2) SEQUENCE DESCRIPTION: SEQ ID NO:188:

Ala	Gly	Asn	Val	Thr	Ser	Ala	Ser	Gly	Pro	His	Arg	Phe	Gly	Ala	Pro	1	5	10	15
Asp	Arg	Gly	Ser	Gln	Arg	Arg	Arg	Arg	His	Pro	Ala	Ala	Ser	Thr	Ala	20	25	30	
Thr	Glu	Arg	Cys	Arg	Phe	Asp	Arg	His	Val	Ala	Arg	Gln	Arg	Cys	Gly	35	40	45	
Phe	Pro	Pro	Ser	Arg	Arg	Gln	Leu	Arg	Arg	Arg	Val	Ser	Arg	Glu	Ala	50	55	60	
Thr	Thr	Arg	Arg	Ser	Gly	Arg	Arg	Asn	His	Arg	Cys	Gly	Trp	His	Pro	65	70	75	80
Gly	Thr	Gly	Ser	His	Thr	Gly	Ala	Val	Arg	Arg	Arg	His	Gln	Glu	Ala	85	90	95	
Arg	Asp	Gln	Ser	Leu	Leu	Leu	Arg	Arg	Gly	Arg	Val	Asp	Leu	Asp		100	105	110	
Gly	Gly	Gly	Arg	Leu	Arg	Arg	Val	Tyr	Arg	Phe	Gln	Gly	Cys	Leu	Val	115	120	125	
Val	Val	Phe	Gly	Gln	His	Leu	Leu	Arg	Pro	Leu	Leu	Ile	Leu	Arg	Val	130	135	140	
His	Arg	Glu	Asn	Leu	Val	Ala	Gly	Arg	Arg	Val	Phe	Arg	Val	Lys	Pro	145	150	155	160
Phe	Glu	Pro	Asp	Tyr	Val	Phe	Ile	Ser	Arg	Met	Phe	Pro	Pro	Ser	Pro	165	170	175	
His	Val	Gln	Leu	Arg	Arg	Ile	Leu	Ser	Leu	Leu	Gly	Gln	Arg	Ser	Ala	180	185	190	
Gln	Phe	Gly	His	Val	Glu	Tyr	Pro	Leu	Pro	Leu	Leu	Ile	Gln	Arg	Ser	195	200	205	
Leu	Ala	Ser	Gly	Ser	Arg	Ile	Ala	Phe	Pro	Val	Val	Lys	Pro	Pro	Gln	210	215	220	
Pro	Leu	Asp	Val	Ala	Leu	Gln	Arg	Gln	Val	Glu	Ser	Val	Pro	Pro	Ile	225	230	235	240
Arg	Lys	Val	Arg	Glu	Arg	Cys	Ala	Leu	Val	Ala	Arg	Phe	Glu	Leu	Pro	245	250	255	
Cys	Arg	Phe	Phe	Glu	Ile	His	Glu	Val	Gly	Phe	Thr	Gly	Arg	Gly	His				

260

265

270

Pro Arg Arg Ile Gly  
275

## (2) INFORMATION FOR SEQ ID NO:186:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 192 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) SEQUENCE DESCRIPTION: SEQ ID NO:186:

```

Arg Val Ala Ala Ser Phe Ile Asp Trp Leu Asp Ser Pro Asp Ser Pro
1      5      10      15
Leu Asp Pro Ser Leu Val Ser Ser Leu Leu Asn Ala Val Ser Cys Gly
20     25     30
Ala Glu Ser Ser Ala Ser Ser Ser Ala Arg Ser Gly Asn Gly Ser Arg
35     40     45
Trp Thr Ser Met Pro Ser Gly Thr Arg Pro Gly Pro Arg Arg Ala Thr
50     55     60
Ser Arg Asp Asp Arg Arg Ser Ala Thr Ser Val Ile Pro Ser Arg Arg
65     70     75     80
Ser Val Ala Pro Arg Ala Glu Phe Gly Thr Arg Leu Ala Ser His Arg
85     90     95
Ala Ser Pro Ser Asn Ala Cys Pro Val Arg Ile Val Thr Ser Ala Ser
100    105    110
Gly Arg Pro Ile Ser Ser Pro Pro Ile Val Arg Ser Arg Ser Cys Val
115    120    125
Asp Lys Asn Gly Arg Arg Cys Ala Ser Gly Tyr Arg Arg Leu Asn Arg
130    135    140
Ala Arg Ser Ser Ser Ile Ala Ala Arg Cys Arg Thr Ile Gly Thr Phe
145    150    155    160
Arg Arg Ser Arg Tyr Ser Ala Ser Met Arg Val Ser Thr Asn Ser Pro
165    170    175
His Val Thr His Gly Val Ala Pro Gly Val Thr Arg Arg Ile Gly Gly
180    185    190

```

## (2) INFORMATION FOR SEQ ID NO:187:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 196 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:

Gln Glu Arg Pro Gln Met Cys Gln Arg Val Ser Glu Ile Gln Pro Arg  
 1 5 10 15  
 Thr Gln Phe Phe Asn Arg Cys Ala Leu Pro His Tyr Trp His Phe Pro  
 20 25 30  
 Ala Val Ala Val Phe Ser Lys His Ala Ser Leu Asp Glu Leu Ala Pro  
 35 40 45  
 Arg Asn Pro Arg Arg Ser Ser Arg Arg Asp Ala Glu Asp Arg Arg Val  
 50 55 60  
 Ile Phe Ala Ala Thr Leu Val Ala Val Asp Pro Pro Leu Arg Gly Ala  
 65 70 75 80  
 Gly Gly Glu Ala Asp Gln Leu His Asp Leu Gly Val Cys Arg Arg Gln  
 85 90 95  
 Ala Gly Arg Val Arg Arg Gly Gln Glu Leu His His Arg His Arg His  
 100 105 110  
 Gln Gly Ala Ala Pro Asp Leu Arg Arg Arg Arg Arg His Arg Arg Val  
 115 120 125  
 Gln Glu His Arg Arg Leu Gln Arg Val Arg Gln Leu Arg Arg Tyr Val  
 130 135 140  
 Gln Thr Ala His His Arg Arg Phe Ala Arg Thr Asp Arg Val Arg His  
 145 150 155 160  
 His Val Arg Gly Pro Ser Asn His Arg Arg Arg Arg Val Tyr Arg Gly  
 165 170 175  
 Arg His Ser Gly Ala Gly Gly Cys Pro Ala Gly Gly Ala Gly Ser Val  
 180 185 190  
 Gly Gly Ser Ala  
 195

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:188:

Val Arg Cys Gly Thr Leu Val Pro Val Pro Met Val Gly Phe Leu Thr

183

1	5	10	15
Ser Thr Asn Ala Pro Ser Ser Leu Phe Ser Ala Tyr Ala Glu Val Asp Lys	25	25	30
Leu Ile Gly Leu Pro Ala Gly Thr Ala Lys Arg Trp Ile Asn Gly Tyr	35	40	45
Glu Arg Gly Gly Lys Asp His Pro Pro Ile Leu Arg Val Thr Pro Gly	50	55	60
Ala Thr Pro Trp Val Thr Trp Gly Glu Phe Val Glu Thr Arg Met Leu	65	70	75
Ala Glu Tyr Arg Asp Arg Arg Lys Val Pro Ile Val Arg Glu Arg Ala	85	90	95
Ala Ile Glu Glu Leu Arg Ala Arg Phe Asn Leu Arg Tyr Pro Leu Ala	100	105	110
His Leu Arg Pro Phe Leu Ser Thr His Glu Arg Asp Leu Thr Met Gly	115	120	125
Gly Glu Glu Ile Gly Leu Pro Asp Ala Glu Val Thr Ile Arg Thr Gly	130	135	140
Glu Ala Leu Leu Gly Asp Ala Arg Trp Leu Ala Ser Leu Val Pro Asn	145	150	155
Ser His Arg Gly Ala Thr Leu Arg Arg Leu Gly Ile Thr Asp Val Ala	160	165	170
Asp Leu Arg Ser Ser Arg Glu Val Ala Arg Arg Gly Pro Gly Arg Val	175	180	185
Pro Asp Gly Ile Asp Val His Leu Leu Pro Phe Pro Asp Leu Ala Asp	190	195	200
Asp Asp Ala Asp Asp Ser Ala Pro His Glu Thr Ala Phe Lys Arg Leu	205	210	215
Leu Thr Asn Asp Gly Ser Asn Gly Glu Ser Gly Glu Ser Ser Glu Ser	220	225	230
Ile Asn Asp Ala Ala Thr Arg Tyr Met Thr Asp Glu Tyr Arg Glu Phe	235	240	245
Pro Thr Arg Asn Gly Ala Glu Arg Ala Leu His Arg Val Val Thr Leu	250	255	260
Leu Ala Ala Gly Arg Pro Val Leu Thr His Cys Phe Ala Gly Lys Asp	265	270	275
Arg Thr Gly Phe Val Val Ala Leu Val Leu Glu Ala Val Gly Leu Asp	280	285	290
Arg Asp Val Ile Val Ala Asp	295	300	
305	310		

(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:



(A) LENGTH: 2672 base pairs  
 (B) TYPE: nuclear acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:189:

CTCTGCGGCA	TTCCGCGACA	GCTGAGCGAG	CCAAAGGGCC	GTTCGAGCAA	GTCTCGAGAG	60
CATTGCGGCA	CGGCGTGGCC	GGCAAGGGTA	AGCAAAATCA	CACGACGGTG	AACGAGCTGT	120
CGAGGCGGT	GAAGCGCTTG	AATGAGGGG	CGGCGAGCTT	CTTCGCGGTG	GTACGCGAGC	180
TGGCGCTATT	CGTCAACCGG	CTACATCAGG	ACGACCAAGA	GTTCGCTGGG	TGACACAAGA	240
ACCTTCGCGA	GTTCACCGAC	AGGTTCAGCC	ACTCGAGTGC	GGAGCTGTGG	AACGCGATCC	300
AGCAATTGCA	CAGCTTCGTC	CGGTCGCGC	GCGCTTCTTT	CGCCAGAGAC	CGGAGGTTGC	360
TGAGCGATGA	CGTCAATAT	CTCGCGAGCG	TACGACCCAC	GTTCGTGGAG	CGGAGTGGT	420
TGGATGGGT	GGAGAGCGTC	CTCGCATAT	TCCGAGCGCT	GGGCGCGAAC	ATTACGAGC	480
TTTACCATCC	GACACAGCGT	GGGTTGGTGT	CGCTTTCGCG	GTTCACGAAT	TTGCGCAACC	540
CGATCGAATT	CATCTCGAGC	TGATTCGAG	CGGATAGCG	GCTCGTATAT	CAAGAGTGGG	600
CGGAGCTCTG	TGGCGAGTAT	CTGGCGGAG	TCTCGAGTGC	GATCAAGTTC	AAGTACTTTC	660
CGTTGGGCTT	GAAGCTGGCC	AGGCGGCTT	CGACAGTGGC	TAAAGAGATC	GCTTACTCGG	720
AGCGCGGCTT	CGACCGGCGC	AACGGGTACA	AGGACCCAC	GGTTCGAGC	ATCTGGGTGC	780
CGGATAGGCT	GTGCTACAG	CGCAACAGG	AGCGCGTTC	GTTCGTGCA	CGCGGATGC	840
AAGGGTTTCA	GGTGGGAGCG	ATACGCGAG	GTTCGTGAC	GGCGGATTC	CTGGCGGAGC	900
TGATGGGTGG	TCCGATATG	GCGCTTCTTT	CGTACGGCT	CGAGAGAGC	CGCGAGCTTC	960
CGAATGGGTA	CGAGAGATAC	CGCTTCTTGC	CGCGATGCG	TTTACAGGCT	CGAGAGTGC	1020
CGATAGGCGC	CGCGCTTCTT	GGCGCGAGC	TATCGCGAG	TCCGCTGCG	CGCGCTTTCG	1080
CGCGATCGT	GTTCGCAAGA	GATCGCGCG	CAGCTTGGG	AACTTGGAG	TGATCGGCG	1140
TCTTGTTCCT	GTTCGCGAGC	CTTCGAGCTT	TCTTGTTCG	GTTCGTATAT	AGCGCGGCGC	1200
GTGGAGAGAT	GGCGGATGCG	CAGCTTCTTC	TACCGCGAG	CAGCGGCTTC	CGCTTGTATG	1260
CGCGATCTGT	CGCATATTCG	TGATCGCGA	CAGAGATTC	GCTCATAGAG	ATCGGCTTGT	1320
TCCGAGAGCG	AGCGGCTGCG	CAGCGTACA	TGAGGATGAC	GCTGCTCTCG	CTCGGCTCTT	1380
TTCGCTCTCT	CTTGTGCTTC	CGAGAGTACG	TCCAGCAAT	GTTCAGCAA	TTCGAGATGC	1440
AATCGGCGGT	GCATATCATC	CGACAGGCGC	TGCTTCTCAT	GCTGCGGATG	CGATCGCGCG	1500
GAGCGATGAT	GGAGCGAGCG	GGAGCGGCGA	AGATGCTGCT	GCTTGGGATC	ATGCTGATCG	1560

CTCCGCGGTT GGAGCCTTC CCTTTGGTG TCGCGGGCA AGCGACTAC TTACCCATC	1620
TGCGGACCG GCTGGCAATC ATGGGCAAGG GATGGGCTG CTGATKATC CACTGTCC	1680
GGGXXGAGT GCAGAGCTG GCGGACATC AGTGGCTCG CGTTGAGCG CTGATCAGCG	1740
TCAACGAGCA GGTGGGCGG TCGATAGGA GCGACTGAT GTGGTCTG CTGACCTAC	1800
AGTTCAATCA CAGCAAAATC ATCGCTACTC CAAGCAAGT CGCACTGACC GAGAGAGTG	1860
GGCGGGGCG GGGGGGGCG GTTGAACCTT CTTGCTACG CGGCAAACT AACTTGGCG	1920
CGCACTGCT GATGACCTT TCGACGCTT AGCGGTGTG ATTGCTGATA GCGAGCGCG	1980
TAGTGTCTC GACCTGATC CCGGGGAT TCTTGGGAA AGCGAGGCT AGTCATCGA	2040
GAGCAGCTT GUTATCCCA TGACTCTGC TT	2072

(2) INFORMATION FOR SEQ ID NO:190:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1923 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(1) SEQUENCE DESCRIPTION: SEQ ID NO:190:

TGACCCGGA GAATCTCTC GTGACGACT TGACATGCA CTGCTCTCG ATGGTCGAGA	60
TGCGGTGCA GACCGAGGAC AAGTAGGCG TCAAGATCC GCGGAGGAC CTGCGGCTC	120
TGCTACCGT CGTGAAGTT GTGCTTACA TCGAGAGCT CGAGGAGGA AACCGGAGG	180
CGGCTAGGC GTTGGCGCG AGATTTGACT CGGAGAAATC CGATCGGCA CGACGAGATC	240
GTGCGTTTC ACCCATAGG CAGCTCGAG AGCCGCTCG TCTCTTGA CCGTCAGGCA	300
GATTTGGGTG TGGCGGCTT CTAGTAGTG TTCCGACAC AGGAGGGAC CCGCGGAAA	360
GGTACTGAT CGCGGAGCA CATAGTCGAT GTCGCTGCG CTGACATTC GCGCGGATC	420
GATTTGGCG GCGCGAATT GCGGATTGG CTGAGGCGG AGCGATCGG GCGCGCGCG	480
CGCGGTGCT GATGTTTGG GCGCGGAT GCGCGGAGG AGGAGGAGG TGGCGCGAT	540
GATGAGGAT AGCGGATCA GAGGAGGAG ATTTCGAG CATACCTCT CGTACCGCTG	600
GCGCGGCTT GGTGATGCG TCGATATCG ATGGGCGGT TTAGGTAAG AGTTTCGCG	660
GGACCGGGG TCAACACGG CGAGTTGTC GCGCGGAGC CGGCGAGTC TCGCGCGGG	720
TGCGCGGCG TCACTGCTG ACATTCGGG TGTGCTGAG GTCGCACTG AAACGACTC	780
AACGCGACG GTTTCGAGG TCACCACTG AGCTGAGC GCGATGCTT GTAGGTTTC	840
GAGCGCGCG AGGTGCGGG TCGGAGCTT TCGCGGCGA GCTTTCGCG TGAAGCGAC	900

UAGGAGTATC	TAGGTTCCGC	CACCGGTGAC	ATCGTGTCTC	GCGAGGTGCT	CGGTCAAGCC	960
GGGATATGAG	CAGGCATCCA	GTCCCAAGTA	GTCTCTCGAG	CTGATGTCCG	CCAGTATGAC	1020
GTGCACGCGA	ACGGGGGCAA	TACCATCGCG	CGGTGGTAGC	CGGCTCAAGA	CCGATATGCT	1080
TTCCACAGCC	GGCTGGCGCA	TGAGATGAG	GGCAGGTTG	AGCGCGCGCA	GGGGCGCTTC	1140
GTGCTCTTGG	TGACAGCTGG	GGATTCGCG	AAAGAGGAGC	CTGGTGTCTG	GTGCGATCA	1200
CGCGGTGTGG	GATCGAGGCT	TTCCCGAAGC	ATTTCGTGGC	TCAACGGGGG	CAGGAGAGGT	1260
TCGTGGGTGG	GGACGAGAAC	CGAGCCTTCC	CGACAGATT	CGACAGGAGT	CGGGCGCGGC	1320
TCAATCTCGA	TGGCGGATC	GGGCTGGTG	ATCTCGAGCT	GGTGGTTCCG	GGCAAGGAGA	1380
AGCGGTCTGC	GAATCGGCT	GGGAATCAGC	AGAGGTCTGC	CGACATCGAT	GGTGTTCGCG	1440
ATGATAGGAA	ATTACCATC	GCAGCTTCCA	TAGGCTGTGC	CTGGCGGCGA	TGTCGGGAGC	1500
ATCGGCTAGC	GTATCGAGC	ATTGTTTGGC	AAATGGCTGA	GGGAGGCTGC	GTTCGGGCTG	1560
ATGGGTGTGG	ATCCCGGCT	GACCGGATGC	GGGCTGTGGC	TCATCGAGAG	TGGCGCTGGT	1620
CGGCACTCCA	CGCGGTGGA	TGTGACGCTG	GTTCGACAC	GGTGGGATGC	GGCCTTGGCG	1680
GAGCGGCTGT	TGGCATCAG	CGATCGGCTC	GAGCACTGGC	TGGACACCCA	TCATCGGAGG	1740
GTGGTGGCTA	TGAAACGGGT	GTCTGTCTAG	CTGAAGGTGA	CGACGCTGAT	GGGACCGGCG	1800
CGGGGGGGGG	GGTGTCTGCG	GTGGGGGGGG	GGCAACGCTG	GTGTGAGGCT	GCATTTCCAT	1860
ACCGCCAGCG	AGGTCAGGAG	GGGGGTGAGT	GGCAACGGTT	CGGCAAGACA	GGCTCAGGTC	1920
ACC						1923

## (2) INFORMATION FOR SEQ ID NO:191:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1955 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) SEQUENCE DESCRIPTION: SEQ ID NO:191:

CTGGGTCTCC	AGTGTGAGGC	GGGATATGAG	GTGGGCAATG	AAATTCGCGG	CGTGGCGGGA	60
CGCGTGGCGA	CGCAATCTGG	AGCACCGGCT	CGGTCATGTC	CGGAAGGTTC	TCAGTGTGCT	120
GGCAATCTGG	GTGTGTGGTT	TCCTGAAGCA	AGGCTGCGCG	TATGGGTGTC	GTGACGCGCA	180
AAGACGCTCA	GTCCAGGAAT	CGGTCGCGCG	CGGGCGGATT	CGGAGGCGCA	TCTGTAGAGC	240
GGGATAGGTT	CAGACGTAA	GGTTGGCTAG	GTGGAAGAGC	GGGCGAGGCG	GGCTGACAGG	300
GGTATGAGCA	GGGAATTAAC	AAAGCGGGGG	ATATTTGTCG	CGGATTGTCA	TAGATGCTGC	360

AGTGGTTGCT	GGTTCTGTGT	TAGCAATTGA	GTGTGGATGT	GTTGAGAGCC	TGGGCTGGAA	420
GGGGACAACG	TGCTTTTGGC	TCTTGGTCCG	CTTTTGGGCT	GGGAGCCGGT	GGTGGGSAAA	460
GGGGTGGAGT	GGGGAATGCT	CGGCGGATTC	TGGGTTGGCG	TCAGCTGGGG	NTTGGCTGTO	540
CGACCCGATG	ATTATGACCA	CTGGGCGGCT	GGGCGGAGGG	ACGGCGGCGA	TGTGATGTTC	600
CAGCGGCGCG	AGGAGGAGGA	CGGAGAGGCC	CGCGGCGATG	AGGAGTGGGA	TGAGTGGGAG	660
GGTGGGAGCG	AGTGGGTGGG	GGGGAACGCT	GAGCGCGGCT	TTGAGTGGCC	ACGGAGTGGC	720
AGGAGGCTGA	TTCGCGATTC	TCCGCGCGCC	GGCTAGGAGA	GGGAGGCGAG	ACTGCTGTTA	780
TTTGACCAAT	GATCGGGGCT	CTCGGTCTTC	CGGCGCGCGG	GTATGACAGC	AGTGAATGTG	840
CATGAGAAGT	TACAGGTATT	AGTCTCAAGT	TCTACAAGGA	GACAGGCAAC	ATGGCAACAC	900
GTTTTATGAC	GGATCGGACG	GGGATGCGGG	ACATGGCGGG	CGCTTTTGGG	GTGACGCGCC	960
AGACGGTGGG	GGGAGAGGCT	CGCGGATGCT	GGGCGTGGCG	GGAAACATCT	TGCGGAGGCG	1020
GGTGAAGTGG	CAAGCGGAGG	GGGAGCTCGC	TAGAC			1055

(2) INFORMATION FOR SEQ ID NO:192:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 359 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

CGGCTGCTTT	GTGGCATAC	TCCGCGGCGG	CGGCTGAGAC	CGGCTGCGCC	GTGGCTGTC	60
TCCGCGCTGA	CCACCGGAGT	CGGGAACCA	TGGAGATCA	CGTTCAGTTC	ATCCAGCTCG	120
CGGCTGCTTT	CACCGAGGGA	CGGCGGCTG	TGGAGAGCG	CGTTCAGTTC	CTGCTATAG	180
CGTTCGCGCC	CGAGCGGAG	GAGTTGTAG	TACTGGCGCA	CGGCTGCTTT	ACCGGAGCG	240
GAGAGTTCA	GGTTCAGGCT	CGGCTGCGG	CGGCGGAGT	AGTTCAGCGG	GAAAGCAGA	300
TCTTCGCGCA	GGTTCGCGG	CGGCGGAGC	AGGAGAGG	CGGCGCGG	ATAGGTCAG	359

(2) INFORMATION FOR SEQ ID NO:193:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 350 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

AACGGGCTCTG TGGGCAAGGC TCTCTAAGG GCTCTGCTTG GTGGCAATGAA GTCGTGGAAS	60
GATGCTCTTT GGGAGATTCG GGGCAGAGCA AACAGGCGCG TACTCTTAGT CCGAGTGGCG	120
CGGAAGGTTC CTCGAATAGC TCGGTAGCGG GAGCCGCAAA CCGGGTCTCC TTCCTAAGC	180
TGGCGAAGC ACTGAGGATT CCGGAGTTCG TTGACGTCCA GAGGATTCG TTGAGTGGC	240
TGATGGTTC GGGGGCTGG CCGGATTCG CCGCGAGCG GGGTGTCTC AACCCAGTGG	300
GTGGCCTTGA AGAGGTGCTC TACGAGCTGT CTCGGATCGA GCACTTCTCC	360

## (2) INFORMATION FOR SEQ ID NO:194:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 679 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xii) SEQUENCE DESCRIPTION: SEQ ID NO:194:

Glu	Gln	Pro	Lys	Gly	Pro	Phe	Gly	Glu	Val	Ile	Glu	Ala	Phe	Ala	Asp	1	10	15
Gly	Leu	Ala	Gly	Lys	Gly	Lys	Gln	Ile	Asn	Thr	Thr	Leu	Asn	Ser	Leu	20	25	30
Ser	Gln	Ala	Leu	Asn	Ala	Leu	Asn	Glu	Gly	Arg	Gly	Asp	Phe	Phe	Ala	35	40	45
Val	Val	Arg	Ser	Leu	Ala	Leu	Phe	Val	Asn	Ala	Leu	Sis	Gln	Asp	Asp	50	55	60
Gln	Gln	Phe	Val	Ala	Leu	Asn	Lys	Asn	Leu	Ala	Glu	Phe	Thr	Asp	Arg	65	70	75
Leu	Thr	His	Ser	Asp	Ala	Asp	Leu	Ser	Asn	Ala	Ile	Gln	Gln	Phe	Asp	80	85	90
Ser	Leu	Leu	Ala	Val	Ala	Arg	Pro	Phe	Phe	Ala	Lys	Asn	Arg	Glu	Val	95	100	105
Leu	Thr	Sis	Arg	Val	Asn	Asn	Leu	Ala	Thr	Val	Thr	Thr	Thr	Leu	Leu	110	115	120
Gln	Pro	Asp	Pro	Leu	Asp	Gly	Leu	Glu	Thr	Val	Leu	His	Ile	Phe	Pro	125	130	135
Thr	Leu	Ala	Ala	Asn	Ile	Asn	Gln	Leu	Tyr	His	Pro	Thr	His	Gly	Gly	140	145	150
Val	Val	Ser	Leu	Ser	Ala	Phe	Thr	Asn	Phe	Ala	Asn	Pro	Met	Gln	Phe	155	160	165

Ile Cys Ser Ser Ile Gln Ala Gly Ser Arg Ileu Gly Tyr Gln Gln Ser  
 180 185 190  
 Ala Gln Leu Cys Ala Gln Tyr Leu Ala Pro Val Leu Asp Ala Ile Lys  
 195 200 205  
 Phe Asn Tyr Phe Pro Phe Gly Leu Asn Val Ala Ser Thr Ala Ser Thr  
 210 215 220  
 Leu Pro Lys Gln Ile Ala Tyr Ser Gln Pro Arg Leu Gln Pro Pro Asn  
 225 230 235 240  
 Gly Tyr Lys Asp Thr Thr Val Pro Gly Ile Trp Val Pro Asp Thr Pro  
 245 250 255  
 Leu Ser His Arg Asn Thr Gln Pro Gly Trp Val Val Ala Pro Gly Met  
 260 265 270  
 Gln Gly Val Gln Val Gly Pro Ile Thr Gln Gly Leu Leu Thr Pro Gln  
 275 280 285  
 Ser Leu Ala Gln Leu Met Gly Gly Pro Asp Ile Ala Pro Pro Ser Ser  
 290 295 300  
 Gly Leu Gln Thr Pro Pro Gly Pro Pro Asn Ala Tyr Asp Gln Tyr Pro  
 305 310 315 320  
 Val Leu Pro Pro Ile Gly Leu Gln Ala Pro Gln Val Pro Ile Pro Pro  
 325 330 335  
 Pro Pro Pro Gly Pro Asp Val Ile Pro Gly Pro Val Pro Pro Val Leu  
 340 345 350  
 Ala Ala Ile Val Phe Pro Arg Asp Arg Pro Ala Ala Ser Gln Asn Phe  
 355 360 365  
 Asp Tyr Met Gly Leu Leu Leu Leu Ser Pro Gly Leu Ala Thr Phe Leu  
 370 375 380  
 Phe Gly Val Ser Ser Ser Pro Ala Arg Gly Thr Met Ala Asp Arg His  
 385 390 395 400  
 Val Leu Ile Pro Ala Ile Thr Gly Leu Ala Leu Ile Ala Ala Phe Val  
 405 410 415  
 Ala His Ser Trp Tyr Arg Thr Gln His Pro Leu Ile Asp Met Arg Leu  
 420 425 430  
 Phe Gln Asn Arg Ala Val Ala Gln Ala Asn Met Thr Met Thr Val Leu  
 435 440 445  
 Ser Leu Gly Leu Phe Gly Ser Phe Leu Leu Leu Pro Ser Tyr Leu Gln  
 450 455 460  
 Gln Val Leu His Gln Ser Pro Met Gln Ser Gly Val His Ile Ile Pro  
 465 470 475 480  
 Gln Gly Leu Gly Ala Met Leu Ala Met Pro Ile Ala Gly Ala Met Met  
 485 490 495  
 Asp Arg Arg Gly Pro Ala Lys Ile Val Leu Val Gly Ile Met Leu Ile  
 500 505 510

Ala Ala Gly Leu Gly Thr Phe Ala Phe Gly Val Ala Arg Gln Ala Asp  
 515 520 525  
 Tyr Leu Pro Ile Leu Pro Thr Gly Leu Ala Ile Met Gly Met Gly Met  
 530 535 540  
 Gly Cys Ser Met Met Pro Leu Ser Gly Ala Ala Val Gln Thr Leu Ala  
 545 550 555 560  
 Pro His Gln Ile Ala Arg Gly Ser Thr Leu Ile Ser Val Asn Gln Gln  
 565 570 575  
 Val Gly Gly Ser Ile Gly Thr Ala Leu Met Ser Val Leu Leu Thr Tyr  
 580 585 590  
 Gln Phe Asn His Ser Glu Ile Ile Ala Thr Ala Lys Lys Val Ala Leu  
 595 600 605  
 Thr Pro Glu Ser Gly Ala Gly Arg Gly Ala Ala Val Asp Pro Ser Ser  
 610 615 620  
 Leu Pro Arg Gln Thr Asn Phe Ala Ala Gln Leu Leu His Asp Leu Ser  
 625 630 635 640  
 His Ala Tyr Ala Val Val Phe Val Ile Ala Thr Ala Leu Val Val Ser  
 645 650 655  
 Thr Leu Ile Pro Ala Ala Phe Leu Pro Lys Gln Gln Ala Ser His Arg  
 660 665 670  
 Arg Ala Pro Leu Leu Ser Ala  
 675

## (2) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 125 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

Thr Pro Glu Lys Ser Phe Val Asp Asp Leu Asp Ile Asp Ser Leu Ser  
 1 5 10 15  
 Met Val Gln Ile Ala Val Gln Thr Glu Asp Lys Tyr Gly Val Lys Ile  
 20 25 30  
 Pro Asp Glu Asp Leu Ala Gly Leu Arg Thr Val Gly Asp Val Val Ala  
 35 40 45  
 Tyr Ile Gln Lys Leu Glu Gln Glu Asn Pro Glu Ala Ala Gln Ala Leu  
 50 55 60  
 Arg Ala Lys Ile Gln Ser Glu Asn Pro Asp Ala Ala Arg Ala Asp Arg  
 65 70 75 80

Cys Val Ser Pro Thr Ser Gln Ala Arg Asp Ala Arg Arg Pro Leu Ala  
 85 90 95  
 Arg Ser Ala Arg Leu Ala Cys Arg Arg Leu Pro Ala Ser Val Pro Thr  
 100 105 110  
 Thr Arg Arg Asp Pro Arg Gln Arg  
 115 120

## (2) INFORMATION FOR SEQ ID NO:196:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 83 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:196:

Leu Ala Cys Gln Cys His Arg Arg Tyr Asp Val Gly Ile Gln Phe Arg  
 1 5 10 15  
 Gly Pro Ala Gly Pro Val Ala Thr Gln Ser Gly Pro Pro Gly Pro Ser  
 20 25 30  
 Ile Ala Gln Gly Arg Gln Val Arg Ala Gln Lys Gly Ala Gly Phe Leu  
 35 40 45  
 Glu Arg Arg Pro Ala Val Ser Gly Ala Leu Pro Pro Asn Asn Ala Ser  
 50 55 60  
 Pro Gly Ile Arg Ser Arg Ala Ala Asp Ser Gln Arg His Leu Leu Ala  
 65 70 75 80  
 Gly Asp Gly Ser Asp Val Thr Val Gly  
 85

## (2) INFORMATION FOR SEQ ID NO:197:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 119 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:197:

Ala Ser Leu Leu Ala Tyr Ser Ala Ala Ala Ala Ser Thr Ala Leu Ala  
 1 5 10 15  
 Val Ala Cys Val Arg Ala Asp His Arg Asp Arg Arg Thr Ile Arg Asp  
 20 25 30



His Leu Ala Met Ile His Leu Ala Glu Leu Val Thr Gln Pro Pro Gly  
 35 40 45

Gly Val Arg Gln Arg Leu His His Leu Gly Ile Ala Val Ala Pro Gln  
 50 55 60

Pro Gln Glu Val Val Val Leu Ala His His Leu Val Thr Gly Thr Gly  
 65 70 75 80

Glu Val Gln Gly Glu Gly Arg His Val Ala Ala Glu Val Val Asp Pro  
 85 90 95

Glu Asn Gln Ile Leu Arg Gln Val Leu Gly Pro Ala Pro His Asp Lys  
 100 105 110

Pro Asp Ala Gly Ile Gly Gln  
 115

## (2) INFORMATION FOR SEQ ID NO:198:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 116 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:198:

Arg Ala Arg Gly His Arg Ser Ser Lys Gly Ser Arg Trp Ser His Glu  
 1 5 10 15

Val Leu Glu Gly Cys Ile Leu Ala Asp Ser Arg Gln Ser Lys Thr Ala  
 20 25 30

Ala Ser Pro Ser Pro Ser Arg Pro Gln Ser Ser Ser Asn Asn Ser Val  
 35 40 45

Pro Gly Ala Pro Asn Arg Val Ser Phe Ala Lys Leu Arg Gln Pro Leu  
 50 55 60

Glu Val Pro Gly Leu Leu Asp Val Gln Thr Asp Ser Phe Glu Trp Leu  
 65 70 75 80

Ile Gly Ser Pro Arg Trp Arg Glu Ser Ala Ala Glu Arg Gly Asp Val  
 85 90 95

Asn Pro Val Gly Gly Leu Glu Glu Val Leu Tyr Glu Leu Ser Pro Ile  
 100 105 110

Glu Asp Phe Ser  
 115

## (2) INFORMATION FOR SEQ ID NO:199:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 511 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:199:

TGCTACGACG	CAATCGCTTT	GGTACAGAT	GTGGATGGCG	GGTTCGCTGC	TGGCGATGCG	60
GTGAAAGCGG	CCGACGTGTT	CGTCGCTTC	GGGGAGACCA	TGAACTGCT	CAAAAGCGTG	120
GTGCGGCGCG	CGATCGATCG	GGTGGCGCAC	GAGCGCAGGT	GCACCCACTG	TGAAACACCA	180
GGGAGTGTTC	CGTTGCGGTT	CGAGTTGCCA	TGAGGGTGGT	GCTGACCGGC	GGGCGCGGCT	240
TCATCGGGTC	GGCGGTGGAT	GGCGGGTAC	GGGCTGGCGG	TCACGACCTG	GTGGCGCTCG	300
ACCGGCTGCT	GGCGCGCGCG	CAGCGGCCAA	ACCGGGTGCT	GCACCGGGCT	TGGCAGGGGG	360
TGACGTGGG	CGACCGACCG	GGCTGGCGCG	CGTTGTGGCG	CGGTGTGGAT	CTGGTGTGTC	420
ACGAGCGCGC	CATGGTGGGT	GCTGGCTGCA	ACGCGCGCGA	CGCACTCGGC	TATCGCGGCG	480
ACAGCGATTT	CGCTACACAG	GTGCTGCTGG	CGAGATGTT	CGCGCGCGCG	GTGCGCGGTT	540
TGGTCTGGG	GTGCTGATG	GTGCTTTAGT	CGCAGCGCGG	CTATGCTGCT	CGCGCGCGTG	600
CGCGCGTGG	CGCGCGCGCG	CGCGCGCGAG	CGGACCTGGA	CAATGGGGTG	TTGAGGAGCG	660
GTGCGCGCGG	GTGCGCGCGG	CGAGTGTGCT	GGCAATGCTG	CGCGAGAGAG	GGCGCGTGGC	720
GGCGCGCGAG	CGGTACCGCG	CGAGCAGGAG	CGCGCAGGAG	CAGTACCGCG	TGGGTGGGTC	780
GGAGCGAGAT	GGCGGTGGCG	TGGTGGGTT	G			811

## (2) INFORMATION FOR SEQ ID NO:200:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 966 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:200:

GTGCGCGGT	GTGGCGGCG	ATGACTTTCG	CGACACCGCG	CGTACTGTC	GGAGATATCG	60
GACTTTGTGT	TGCGGTGCG	GGGATAGAGG	AGTTGTGCGG	GTGCTGAGG	GTGACCGTTT	120
GTGCGGACCG	CGAACCGGCG	CTTTCAAGGT	AGCTGTGCGG	TGACACAGGT	CGGAGCGGTA	180
AGTTCAGGCT	CGAATATCGG	GTGGAATTCG	GGCGTGAAGT	TGCGTTCGCG	GACATTCAGG	240
CGTACTTCAG	TTAGTTCGCG	GGCAATTTGA	CGGTTTCGAT	CGGTTTCGCG	CTGCTGACCG	300
TGCGGCTCAA	GGTGTACAGG	GGTACCGCGG	AGCAGCAGAT	CGGTTTCGCG	CGGTTTCGCG	360

CCAGAGACAA CCGACGATC CGGTACAAGC GCGTCTGCG GCGGTGTGC GAGGTGCTG	420
ACTACCGCGA TGTTCGCCG GGTACGAGT CCGGCGACG GCAATGGTG GCGATACCG	480
ACGACGACAT CCGACGCTG CCGAAGAAC GCAGCGCGA GATCGAGTG TTGGAGTTG	540
TCCCGCGCG CCGACGGAC CCGATGATG TCGACCGAC CTACTTTTC GAGCGTATT	600
CGAGTCGTC GAATCTGAT GCGTGTCTG CTAGGACCT GCGCGAGAC GCGCGATG	660
CGTCTGTGA TCGCGTACC GCGGTGAT GCGAGAAAA TACGAGCGG TATCCACAAT	720
TGCGCTCGA GCTCGGTAC GAGAACGCT AGACGATCG AGACATTCG CCGCTAGAT	780
GCGCGCTAT AGAAGCGCT CTGCGGATC ATCAAGCGA AATACGCTT ACTCATGCA	840
TGCGCGTGC TCACCGGAT GCGATTTTT GCGCGCTGC ACCGCTGCC GCGCGATTC	900
AAGTGGGAT GCATCGGAC GTTCGCGA AACCGTTG GCGCGTGG CTGATCGCT	960
CATCT	966

## (2) INFORMATION FOR SEQ ID NO:201:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2367 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (2) SEQUENCE DESCRIPTION: SEQ ID NO:201:

GCGACGCGG GCAATACCG CAGCGGAC GGTACCGG GTTGGCGG TTGCGCGCT	60
TGCGCGGCT CCGCGCGCG CCGCGATCG AGTTCTATC GCGAAGTA CTGCGTTTC	120
GCGCGAGCG GCGTTGCG CCGTACCG CAGCGCGCG GACTCAGG GCGCGCGA	180
CTCGCGGCT GCGCGGTC CCGCGTTC GATCAACAT GCGCTGGCG CAGCGCTTC	240
CAGCGCGCG ACCGCTCG CCGCGCGCG CAGCGCGG CAGCTGTC CCGCGCGAC	300
CATCAGGAC TACCGCGCG ACCCGCGA CAGCGCGC CGCTCTTC TGAACGTCG	360
CGTTCGCG ACCGCGCG TTGCGCGA CCGCGCGG AGCGCGCG CCGCATCG	420
CAGCGCGCG GCGTTGCG CCGTTCGCG GTTTCGCG CAACACCG CCGCGCGCG	480
GTTTCGCG CCGCGCGCG GTTCGCGG GCGCGCGG GCGAGCGG CTGCGCGCT	540
TGCGCGCTC ACCCGCTC CCGCGCGA CATCGCTTC TCGCTGCG TCGCGCTCT	600
CAGCGCGG GATCGCGG TTGCGCGG TTGCGCGG CCGCGCGG GCGCGCTAC	660
CAGCGTAC ACCCGCGA CCGCGCGA CCGTTCGCG ATCAGCGAT AGCAACCGG	720
GCGCGCGC ATTGCGCG CCGCGCGG CCGCGCGG GCGCGCGG GCGCGCTAG	780

CAGCCCGCTT	ACCAGCGAAA	CCCGCGCTAC	CACCGGTAGA	GCTGGCAGTG	GCGATGTATA	840
CGAAGCGCGC	GCTCTCGAGG	CCCGCGCTAC	CACCGCCACT	GCGGCGCGCT	ACACCGTCGG	860
ACCGGTTGCC	ACCATCAGCG	CGAAGCGCGC	TGGCAATGTC	GCGCTCGCGG	ACTCGCGCGT	940
CCCGCGCGCT	GCGCGCGCGG	CCACCGCGCG	CGCGCGTACC	GCGCTGACTG	CCCGACCGCG	1020
CGGTGGCGTT	GCGCGAGCGT	GCGCTCGCGG	TGGCAGCGTC	GCGCTCGCGT	GCGCGGTCGG	1080
GCTGGCGCGG	AGTGGCGATG	CGCGCGCGTG	CGCGGTCGCG	GCGCGTTTGA	TGCGCGATCG	1140
CGGACACATC	TGGCGGCGTG	TGGCGGTCG	TGGCGCGCGG	GCGCGCGTGG	GCGATTGACG	1200
CGTTTGCGCG	GCGCGAGCGG	GCGCGCGCGG	TACCGCGCGG	GCGCGCGTGG	GCGAACCGCG	1260
CGCGGTGGCG	GCGGTTACGG	CGCGCACCGG	CGATGCGTGC	GCGCGCGTGG	GTCGCGCGCG	1320
CACCGCGCGT	GCGCGCGTGG	CGCGCACCGG	ACCGCGCGGT	CGCGCGCGCG	CGCGCGCGCG	1380
CGCGCGTACG	ACCGCGCGCG	CGCTTGGCGG	CGTTGGCGAT	CGACCGCGCG	GCGCGTGGCG	1440
TGGCGCGCGT	TTCACCGGAG	CGCGCGACCG	CGCGGTTGGG	ACCGTTCGCG	ACCGCGGACG	1500
GCGCGCGCGG	GCGAGGTCGG	CGCGGTCGCG	TGGCGTGGCG	GCGGTTGGCG	ATCGACCGCG	1560
GCGCGCGGAG	GCGCGCGTGG	GCGCGGTCG	CGCGCGCGCG	CGACCGCGCG	TGACACCGCG	1620
CTTCAGTGTG	GCGATACCGG	CGCGCGCGCG	CGATCGCGCG	CTGCGCGGAG	TGCTCGTGGG	1680
ACCGTGGCGG	CTGTCGCGTG	AGCGCGTGGT	ACGCGCGGAG	ATGCGCGCGG	AACACCGCGG	1740
CAATCGCGCG	CGACAGTTCG	TGGCGCGCGG	CGCGCGCGCG	TTCGCTCGCG	GCGATCGCGG	1800
CGCGCGCGT	AGCGCGCGTG	ACGTCGCGAG	CAATAGTTCG	TAAATCGGAG	GCGCGCGTGG	1860
CGCGCGCGTG	CGCGCGCGCG	ATCGCGCGCG	ACCGCTCGCG	CGCTCGCGCG	CGCGATTCG	1920
CGCGCGCGTG	TGGCGCGCGG	CGCGCGCGCG	TTCGCTCGCG	GCGTGGCGCG	CGTACGCGTG	1980
GCGCGCGCGG	CGCTCGCTCG	GATTCGCGCG	GCGCGCGCGG	TTCGCGCGCG	ATTCGCGCGG	2040
GCGCGCGCGG	GCGCGCGTGG	GTCGCGCGTG	GCGCGCGCGG	GCGCGCGCGG	GTCGCGCGCG	2100
AGATGATAG	GCGCGCGCGG	ATTCGCTTGG	TGCTGAGTTC	CTCGCGCGCG	ACTGAGCGCG	2160
CTTCGCGCGG	TGTCGCGTTC	AACCGCGCGG	AGCGCTTTCG	CGCGCGTTCG	CTTCGCGCGG	2220
GCGCGCGTGG	GCGTACCGCG	ATCGCGTTCG	CGATCGCGCG	GCGCGTTCG	TGCGCGCGCG	2280
CTTCGCGCGG	CTTCGCGTTC	CGATGAGTTC	TTCGAGCGCG	GCGTACCGCG	GAGTACCGCG	2340
GCGCGCGTTC	GATTCGCGCG	CGCGCGCG				2367

(2) INFORMATION FOR SEQ ID NO:202:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 376 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

Gln	Pro	Ala	Gly	Ala	Thr	Ile	Ala	Ala	Ser	Ser	Pro	Cys	Ala	Thr	Val
1			5						10					15	
Gly	Ala	Gly	Gly	Gly	Thr	Gly	Ser	Pro	Val	Thr	Thr	Glu	Thr	Ala	Ala
		20						25					30		
Thr	Thr	Gly	Arg	Gly	Gly	Ser	Gly	Asp	Val	Tyr	Glu	Ser	Ala	Ala	Ser
		35					40					45			
Gly	Ala	Ala	Ala	Thr	Thr	Pro	Thr	Ala	Gly	Gly	Tyr	Thr	Val	Gly	Pro
		50				55					60				
Val	Ala	Thr	Ile	Thr	Ala	Lys	Gly	Ala	Arg	Asn	Val	Ala	Leu	Arg	Asp
65				70					75					80	
Ser	Ala	Val	Ala	Ala	Val	Ala	Ala	Ala	Ala	Thr	Gly	Ser	Gly	Gly	Thr
			85						90					95	
Ala	Val	Thr	Thr	Gly	Thr	Ala	Gly	Gly	Leu	Ala	Arg	Ala	Cys	Arg	Arg
		100					105						110		
Gly	Gly	Thr	Val	Ala	Ala	Gly	Ala	Thr	Gly	Arg	Arg	Ala	Gly	Ser	Ala
		115					120					125			
Met	Ala	Ala	Arg	Ala	Ala	Val	Ala	Ala	Gly	Leu	Ile	Thr	Asp	Ala	Gly
		130				135					140				
His	Ile	Cys	Arg	Ala	Val	Pro	Gly	Ala	Gly	Arg	Gly	Ala	Gly	Arg	Gly
145			150						155					160	
Ile	Asp	Pro	Val	Cys	Pro	Gly	Glu	Ala	Gly	Ala	Ala	Gly	Thr	Thr	Gly
			165						170					175	
Ala	Ala	Met	Ala	Glu	Glu	Pro	Gly	Val	Ala	Ala	Val	Thr	Ala	Arg	Thr
		180						185						190	
Pro	Asp	Ala	Cys	Gly	His	Ala	Gly	Ala	Ala	Asp	Thr	His	Val	Ala	Ala
		195					200					205			
Val	Ala	Pro	Gln	Pro	Pro	Pro	Val	Pro	Thr	Gly	Thr	Ala	Gly	Arg	Ala
		210				215					220				
Gly	Thr	Thr	Gly	Pro	Ala	Val	Ala	Ala	Val	Ala	Asp	Gln	Pro	Gly	Arg
225				230					235					240	
Ala	Ser	Ala	Ala	Ala	Gly	Leu	Thr	Gln	Pro	Ala	Ser	Arg	Ala	Val	Ala
			245					250						255	
Thr	Val	Ala	Lys	Gln	Gln	Pro	Ala	Gly	Arg	Ala	Arg	Leu	Pro	Gly	Cys
		260					265						270		
Arg	Pro	Val	Gly	Ala	Val	Ser	Asp	Gln	Arg	Ala	Pro	Gln	Lys	Arg	Leu
		275				280						285			
Gly	Gly	Arg	Ile	His	Arg	Thr	Gln	Gln	Thr	Pro	Leu	Asn	Ser	Gly	Phe
		290				295					300				

Ser Ala Gly Ile Pro Thr Arg Gly Arg Ser Gln Arg Leu His Lys Leu  
 305 310 315 320  
 Leu Val Lys Arg Cys His Leu Tyr Ala Gln Arg Leu Ile Leu Pro Ser  
 325 330 335  
 Met Gly Pro Gln Gln Pro Arg Asn Arg Arg Arg His Phe Ile Gly Ser  
 340 345 350  
 Arg Ser His His Phe Arg Arg Arg Asp Arg Arg Gly Arg Ile Ser Arg  
 355 360 365  
 Ala His Leu Arg Thr Asn Ser Arg  
 370 375

## (x) INFORMATION FOR SEQ ID NO:203:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2852 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

GGCAAAACG CCGCGGCAT CCGGCCAC GAGGCGCTT ACACACGAT CTGGGCCAG 60  
 GACCTCGCG CAGTATTGG CTACATGCC GGGCTTGG CGGCGTCTC GCGTTGACA 120  
 CGCTTCGAC AGGCGTGGC GAGCTGGG GGGGGCGTG CGCTGGTCA CCGGCGCGG 180  
 GCTCAGTCA CCACGGGGT CTTCGGCAC CTGGCTTGG CGACGTCGC CGAGGCGAC 240  
 CTCGCCAGC GATATCTCG GACTTCAAT CTGGCTTGG CGACATCGG CACGCGAAC 300  
 ATGCGACGC GCAACATCG CAGCTCAGC ATGGGTTT GAAAGTCGG TCTGGTTT 360  
 AGCGCAGCG TGACCAAT CCGTTTCAG AACACGGCA GACCAACAT CCGGTTTGG 420  
 AACCGGCGA GCAACACAT CCGGTTTGG AATACGGAG AGCGCAACG AGATATCGG 480  
 CTCACGGTA CGGTTTATT GCGTTTCGG GCGTGAAT CGGCAACCG CAACTCGAT 540  
 CTGTTCAAT CGGCGACCG AAGCTGCGC ATCGGACAT CGGTACCG GACTTGGCG 600  
 ATTCGCAAT CGGCGACCG CTACACGCG GCTTTGCGA ACTCGGCGA CGGCAACCG 660  
 GGTCTTCA ACTCGGAT AGCGACGCG GCGTGGCGA AGCGCGGCA CTACACGCG 720  
 GGTACTTCA AGCGGCGA CAGCAATAC GCGGCTTCA ACTCGGCGA CTACACGCG 780  
 GGTACTTCA AGCGGCGA CTACACGCG GCTTTGCGA ACTCGGCGA TGTACGCG 840  
 GCGGCTTCA TTACTGCGA CTACACGCG GCGTCTTGT GCGGCGGCG CAGCAACCG 900  
 CTGATTTCG GAGCGCGCG CTCTTCGCG TGACGCGT CCGGCTTGT GCGTCTTGT 960

AACAGCGGTC	CGGTATGDC	CTCGGGCTTC	CTGAAGTCG	GTGGCAACA	TTTGCGCTTC	1020
TTCACTCTT	CGTCGGGGG	CATCGGTAC	TCCGGGCTG	CAACCGGG	CGTCTGGA	1080
TGGGCGTGA	TCACCTGGG	CAACACGTA	TGGAATTTG	TCAACATAG	CGTCTGACC	1140
ATCACAGCG	CGGCTTGAT	CTCGGGCTTC	TTCAACGCG	GAAGCAACT	GTGCGGATT	1200
TTGCTGGGC	GACCGGCTT	CAATCTGGG	CTGCAAAAC	GGGGCGTGT	GAACATTCG	1260
GGCAACGCA	ACATCGGAA	TTACAACAT	CTGGCGAGG	CAACCGTGG	TGACTTCAC	1320
ATGCTTGCA	GGGCAACTT	CGGCAACAA	AACATCTTG	GCAGCGGCA	CGTGGGAGC	1380
TTCAATATG	GCATCGGAA	CATCGGAGT	TTCAATCTG	GTTCGGGAG	CGTGGGAGC	1440
TACAACATG	GATCGGAAA	CGTCGGGAT	TACAACATG	GTTCGGGAA	CGTGGGAGC	1500
TACACGCTG	GCTTCGGAA	CGCGGGGAG	TTCAACCAAG	GCTTCGGAA	CGCGGGGAG	1560
AACACATCG	GCTTCGGAA	CACCGGAGC	AACACATCG	GCATCGGAT	GTTCGGGAG	1620
AACGCGGAG	GCTTCGATAT	TGCTAGCGG	TGCAACTGG	GCACCGGAA	CGCGGGGAG	1680
TTCAATCTG	GCACCAATG	CGTTCGATG	TTCAACGCG	GCACCGGAA	CGTTCGATG	1740
GCACACTCG	GCACCGGAA	CTGGGATG	GGGACCGG	GTACCGGAA	TACCGGATG	1800
CTCAATCTG	GCACCTACAA	CACCGGATG	CTCAACCGG	GCACCTTGA	CACCGGATG	1860
TACACAGCG	GCACCTACAA	CACCGGCGG	TTCAACCTG	GTACCGGAA	CACCGGAGC	1920
TTCAACCTG	GTACCGGAA	TACCGGAGC	TATACCGG	GTACCGGAA	CACCGGCTG	1980
TTCAATCTG	GCACCTTGA	TACCGGCGT	TTCAACCGG	CGGCTTGA	CAATCGGCTG	2040
TTGCTGGCG	GCATACGCA	GGCGGAGT	GCATCGATG	CTCGGCGC	CACTCGATG	2100
ATCGGATGA	ACGCGGAGT	GCTCATTCAG	GTACCGGAG	TATGAGCTT	CGCGGCGAG	2160
ATGCTCGAG	TCACCGGAG	CTCGGCGCT	TTCGCGGAA	CGCTTATCT	CGCGGCTTG	2220
TTCTCTCTG	GGCGGCTGA	TCTCGGCGA	TGCGGCTGA	CGCTTCTGAG	GATCGGCTG	2280
ACGCTCGCG	GACCGGAGT	GACCGTCCG	ATCGGCTTG	TGCGGCTTG	GGAGCGGCG	2340
ACGATTCAT	TCTCGGAGT	CGATCGGCG	CGCGGCTTG	GATTCGAG	CACGAGCGG	2400
TGCTCGGCT	TCTCGGAGT	GGCGGCGG	GCGGCTTG	GCTTCGAA	CGTTCGCGG	2460
GGGCTTCAG	GCTTCGAA	CAGGCTTG	GCAGCGG	TAGCGGATG	GGGCTTCAG	2520
AACCTCGGCT	CGTTCGAGT	AGGCTGGCG	AACCTCGG	GCAGCGG	CGGCTTCG	2580
AACCGAGT	CGTTCGAGT	CTCGGCGG	GCAGCTTG	CGCGGCTGA	CACGAGCGG	2640
ACGAGCTCT	CGCGGCTTG	CGCGGCTTG	ACCGGAGT	TTTCGAGG	GGGCTTCG	2700
AACCTGGCG	AGTTCGAGT	CGCGGCGG	TGCTCGGAG	TTTCGAGG	GTTAGATAG	2760
GTTTCGAGG	TATATCTCG	GTTTCGAGG	AGTTCGAG	AGTTCGAG	CGCGGAGG	2820

GTAAGCGGAT AACCGGATG GGGGCTGTC AT

2652

(2) INFORMATION FOR SEQ ID NO:204:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 943 amino acids  
 (B) TYPE: amino acid  
 (C) STANDARDNESS:  
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

Gly Gln Asn Ala Pro Ala Ile Ala Ala Thr Gln Ala Ala Tyr Asp Gln  
 1 5 10 15  
 Met Trp Ala Gln Asp Val Ala Ala Met Phe Gly Tyr His Ala Gly Ala  
 20 25 30  
 Ser Ala Ala Val Ser Ala Leu Thr Pro Phe Gly Gln Ala Leu Pro Thr  
 35 40 45  
 Val Ala Gly Gly Gly Ala Leu Val Ser Ala Ala Ala Gln Val Thr  
 50 55 60  
 Thr Arg Val Phe Arg Asn Leu Gly Leu Ala Asn Val Arg Gln Gly Asn  
 65 70 75 80  
 Val Arg Asn Gly Asn Val Arg Asn Phe Asn Leu Gly Ser Ala Asn Ile  
 85 90 95  
 Gly Asn Gly Asn Ile Gly Ser Gly Asn Ile Gly Ser Ser Asn Ile Gly  
 100 105 110  
 Phe Gly Asn Val Gly Pro Gly Leu Thr Ala Ala Leu Asn Asn Ile Gly  
 115 120  
 Phe Gly Asn Thr Gly Ser Asn Asn Ile Gly Phe Gly Asn Thr Gly Ser  
 130 135 140  
 Asn Asn Ile Gly Phe Gly Asn Thr Gly Asp Gly Asn Arg Gly Ile Gly  
 145 150 155 160  
 Leu Thr Gly Ser Gly Leu Leu Gly Phe Gly Gly Leu Asn Ser Gly Thr  
 165 170 175  
 Gly Asn Ile Gly Leu Phe Asn Ser Gly Thr Gly Asn Val Gly Ile Gly  
 180 185 190  
 Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn Ser Gly Asn Ser Tyr  
 195 200 205  
 Asn Thr Gly Phe Gly Asn Ser Gly Asp Ala Asn Thr Gly Phe Phe Asn  
 210 215 220  
 Ser Gly Ile Ala Asn Thr Gly Val Gly Asn Ala Gly Asn Tyr Asn Thr  
 225 230 235 240  
 Gly Ser Tyr Asn Pro Gly Asn Ser Asn Thr Gly Gly Phe Asn Met Gly



285					290					295					
Gln	Tyr	Asn	Thr	Gly	Tyr	Leu	Asn	Ser	Gly	Asn	Tyr	Asn	Thr	Gly	Leu
				286					285					276	
Ala	Asn	Ser	Gly	Asn	Val	Asn	Thr	Gly	Ala	Phe	Ile	Thr	Gly	Asn	Phe
				275					280					285	
Asn	Asn	Gly	Phe	Leu	Trp	Arg	Gly	Asp	His	Gln	Gly	Leu	Ile	Phe	Gly
				290					295					300	
Ser	Pro	Gly	Phe	Phe	Asn	Ser	Thr	Ser	Ala	Pro	Ser	Ser	Gly	Phe	Phe
				305					310					315	
Asn	Ser	Gly	Ala	Gly	Ser	Ala	Ser	Gly	Phe	Leu	Asn	Ser	Gly	Ala	Asn
				325					330					335	
Asn	Ser	Gly	Phe	Phe	Asn	Ser	Ser	Ser	Gly	Ala	Ile	Gly	Asn	Ser	Gly
				340					345					350	
Leu	Ala	Asn	Ala	Gly	Val	Leu	Val	Ser	Gly	Val	Ile	Asn	Ser	Gly	Asn
				355					360					365	
Thr	Val	Ser	Gly	Leu	Phe	Asn	Met	Ser	Leu	Val	Ala	Ile	Thr	Thr	Pro
				370					375					380	
Ala	Leu	Ile	Ser	Gly	Phe	Phe	Asn	Thr	Gly	Ser	Asn	Met	Ser	Gly	Phe
				385					390					395	
Phe	Gly	Gly	Pro	Pro	Val	Phe	Asn	Leu	Gly	Leu	Ala	Asn	Arg	Gly	Val
				405					410					415	
Val	Asn	Ile	Leu	Gly	Asn	Ala	Asn	Ile	Gly	Asn	Tyr	Asn	Ile	Leu	Gly
				420					425					430	
Ser	Gly	Asn	Val	Gly	Asp	Phe	Asn	Ile	Leu	Gly	Ser	Gly	Asn	Leu	Gly
				435					440					445	
Ser	Gln	Asn	Ile	Leu	Gly	Ser	Gly	Asn	Val	Gly	Ser	Phe	Asn	Ile	Gly
				450					455					460	
Ser	Gly	Asn	Ile	Gly	Val	Phe	Asn	Val	Gly	Ser	Gly	Ser	Leu	Gly	Asn
				465					470					475	
Tyr	Asn	Ile	Gly	Ser	Gly	Asn	Leu	Gly	Ile	Tyr	Asn	Ile	Gly	Phe	Gly
				485					490					495	
Asn	Val	Gly	Asp	Tyr	Asn	Val	Gly	Phe	Gly	Asn	Ala	Gly	Asp	Phe	Asn
				500					505					510	
Gln	Gly	Phe	Ala	Asn	Thr	Gly	Asn	Asn	Asn	Ile	Gly	Phe	Ala	Asn	Thr
				515					520					525	
Gly	Asn	Asn	Asn	Ile	Gly	Ile	Gly	Leu	Ser	Gly	Asp	Asn	Gln	Gln	Gly
				530					535					540	
Phe	Asn	Ile	Ala	Ser	Gly	Trp	Asn	Ser	Gly	Thr	Gly	Asn	Ser	Gly	Leu
				545					550					555	
Phe	Asn	Ser	Gly	Thr	Asn	Asn	Val	Gly	Ile	Phe	Asn	Ala	Gly	Thr	Gly
				565					570					575	

Asn Val Gly Ile Ala Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn  
 580 585 590  
 Pro Gly Thr Asp Asn Thr Gly Ile Leu Asn Ala Gly Ser Tyr Asn Thr  
 595 600 605  
 Gly Ile Leu Asn Ala Gly Asp Phe Asn Thr Gly Phe Tyr Asn Thr Gly  
 610 615 620  
 Ser Tyr Asn Thr Gly Gly Phe Asn Val Gly Asn Thr Asn Thr Gly Asn  
 625 630 635 640  
 Phe Asn Val Gly Asp Thr Asn Thr Gly Ser Tyr Asn Pro Gly Asp Thr  
 645 650 655  
 Asn Thr Gly Phe Phe Asn Pro Gly Asn Val Asn Thr Gly Ala Phe Asn  
 660 665 670  
 Thr Gly Asp Phe Asn Asn Gly Phe Leu Val Ala Gly Asp Asn Glu Gly  
 675 680 685  
 Gln Ile Ala Ile Asp Leu Ser Val Thr Thr Pro Phe Ile Pro Ile Asn  
 690 695 700  
 Glu Gln Met Val Ile Asp Val His Asn Val Met Thr Phe Gly Gly Asn  
 705 710 715 720  
 Met Ile Thr Val Thr Glu Ala Ser Thr Val Phe Pro Gln Thr Phe Tyr  
 725 730 735  
 Leu Ser Gly Leu Phe Phe Phe Gly Pro Val Asn Leu Ser Ala Ser Thr  
 740 745 750  
 Leu Thr Val Pro Thr Ile Thr Leu Thr Ile Gly Gly Pro Thr Val Thr  
 755 760 765  
 Val Pro Ile Ser Ile Val Gly Ala Leu Glu Ser Arg Thr Ile Thr Phe  
 770 775 780  
 Leu Lys Ile Asp Pro Ala Pro Gly Ile Gly Asn Ser Thr Thr Asn Pro  
 785 790 795 800  
 Ser Ser Gly Phe Phe Asn Ser Gly Thr Gly Gly Thr Ser Gly Phe Gln  
 805 810 815  
 Asn Val Gly Gly Gly Ser Ser Gly Val Trp Asn Ser Gly Leu Ser Ser  
 820 825 830  
 Ala Ile Gly Asn Ser Gly Phe Gln Asn Leu Gly Ser Leu Gln Ser Gly  
 835 840 845  
 Trp Ala Asn Leu Gly Asn Ser Val Ser Gly Phe Phe Asn Thr Ser Thr  
 850 855 860  
 Val Asn Leu Ser Thr Pro Ala Asn Val Ser Gly Leu Asn Asn Ile Gly  
 865 870 875 880  
 Thr Asn Leu Ser Gly Val Phe Arg Gly Pro Thr Gly Thr Ile Phe Asn  
 885 890 895  
 Ala Gly Leu Ala Asn Leu Gly Gln Leu Asn Ile Gly Ser Ala Ser Cys  
 900 905 910

Arg Ile Arg His Glu Leu Asp Thr Val Ser Thr Ile Ile Ser Ala Phe  
915 928 925

Cys Gly Ser Ala Ser Asp Glu Ser Asn Pro Gly Ser Val Ser Glu  
930 935 940

(2) INFORMATION FOR SEQ ID NO:205:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 53 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:205:

GGATGCATAT GGCCATCAT CATCATCATC ACCTGATCGA CATCATCGGG ACC 33

(2) INFORMATION FOR SEQ ID NO:206:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CTGTATATCA GGCTTCGTT GCGCCGACCT CATCTTGAC GA 42

(2) INFORMATION FOR SEQ ID NO:207:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:207:

GGATGCTGCA GCTTCGAGG CACCGACCGG T 31

(2) INFORMATION FOR SEQ ID NO:208:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 31 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:208:

CYCTGKATTC AGCCTTGGA ATCCTGCGA T

31

(2) INFORMATION FOR SEQ ID NO:209:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:209:

GAATCCAGG CTGAGATGAA GACCATGCC GGT

33

(2) INFORMATION FOR SEQ ID NO:210:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:210:

GGATATCTGC AGATTTCAGC TTAAAGCCG ATTTCCGA

36

(2) INFORMATION FOR SEQ ID NO:211:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:211:

CGCATGCGA GCGAGTGC CAGACGGCC

30

(2) INFORMATION FOR SEQ ID NO:212:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:212:

CTTCATGGAA TTCTGAGGC GGTAAAGTCC GCTGCTG

37

(2) INFORMATION FOR SEQ ID NO:213:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1676 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:213:

TGGGGAATGG GAGCGGCGCT GTAGCGCGCG ATTAAAGCGG GCGGGTGTTG TGGTTACGGG	60
CAGCGTGACC GGTACACTTG CAGCGGCGCT AAGCGCGGCT CCTTTGGTGT TCTTCGGCTG	120
CTTCTCGCCG ACGTTCGCGG GCTTTCGCGG TCAAGTCTCA AATCGCGGCG TCGCTTAAAG	180
GTTCGCAATT AGTGTCTTAC GGCACCTCGA CCCCCAAAAA CTTGATTTCG GTGATGGTTC	240
ACGTAGTGGG CGATCGGCGT CATACACGGT TTTTCGCGCT TTGACGTTGG ACTCCAGCTT	300
CTTTAAATAG GAGCTCTTGT TCCAACTCG AACAACACTC AACCCATCTT CBTCTATCTC	360
TTTGTATTTA TAAGGGAATT TGGCGATTTC GCGCTATTGG TTAAGAAATG AGCTGATTTA	420
ACAAAAATTT AATCGCAATT TTAACAAAT ATTACGTTT ACATTTTCA GTCGCATTTT	480
TGGGGGAAAT GTGCGCGGAA CCGCTATTTC TTTATTTTTC TAATACATTT CAAATATGTA	540
TGCGCTCATG AATTAACTCT TAGAAAAACT CATCGAGCAT CAATGAAAC TGCATTTTAT	600
TCATKTCAGG ATTATCAATA CACATTTTTT GAAAAAGCGG TTTCTTTTAT GAGGAGAGAA	660
AGTCACCGAG GCGGTTCCAT AGGATGGGGA GATCTGTGTA TCGGTTCGCG ATTCCGACTC	720
GTCCGACATC AATACAGCTT ATTAATTTCC CTTGCTGCAA AATAGGTTTA TCAATGAGA	780
AATCAGCATG AGTCAGGACT GAATCCGCTG AGAATGGGGA AAGTTTATGG ATTCTTTTCC	840
AGGCTTGTTG AAGACGGGAG CCGTTACGCT CGTCATCAAT ATCATTCGCA TCAACCAAAAC	900
CGTTATTGAT TCGTATTTCG GCGTAGCGGA GAGGAAATAC GCGATCGGCT TTAAAGGAC	960
AATTACAAAC ACGAATGGBA TCGAACCGGC GCGGAGACAC TCGCAGCGCA TCAACATAT	1020
TTTACGCTGA ATCAGGATAT TTTTCTAATA CCGGAAATGC TGTTTTCCCG GCGATCGGAG	1080
TGGTGAGTAA CCAATGATCA TCAGGAGTAC GGTAAAAATG CTTGATGCTC GGAAGAAGCA	1140

TAAATTCGGT	CAGCGAGTTT	AGCTGAGCA	TCTGATCTST	AACATCATTC	GCAGGCGTAG	1200
CTTTGGCATG	TTTCAGAAAC	AACCTTGGCG	CATCGGCTT	CCCTAGCAAT	CGATAGATTG	1260
TCCGACCTGA	TTGCTCGACA	TTATCGCGAG	CCCATTTTAT	CCCATATARA	TCAGCATCCA	1320
TCTTGGAAAT	TAACTGCGCG	CTAGAGGAAC	AGCTTTCGCG	TTGAATATCG	CTCATTAACAC	1380
CCCTGTATAT	ACTGTTTATG	TAAGCAGACA	GTTTATTGCT	TCATGAGCAA	ATCTGCTTAA	1440
CGTGAGTTTT	CGTTCCACTG	AGCTTCGAGC	CCGTTGAAA	ACATCAAGAG	ATCTTCTTGA	1500
GATCTTTTST	TTCTGCGCTT	AATCTGCTTC	TTGCAACAA	AAAAACCACT	GCTATCAGCG	1560
GTGTTTGT	TGCCGAGTCA	AGAGCTAGCA	ACTCTTTTTC	CGAGGTTAAC	TGCTTCAGC	1620
AGAGCGGAGA	TACCAAAATC	TGCTCTCTA	GCTAGGCTT	AGTTAGGCTA	CCACTTCAGG	1680
AACCTGTTAG	CAGGCGCTAC	ATACCTCGCT	CTGCTATGCC	TGTTACCACT	GGCTCTCTGC	1740
AGTGGCGATA	ACTGCTGTCT	TACCGGCTTG	GAUTCAAGAC	GATACCTTAC	GGATAAGGCG	1800
CAGCGGTCTG	GTTCAGCGCG	GGTTTCTGCT	ACACAGGCTA	GCTTGGAGCG	AACGACCTAC	1860
ACCGAAGCTA	GATACCTTCA	CGCTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGAGCA	1920
AAGCGGAGCA	GGTATCGCTT	AACCGGCACT	CTCGGAGACG	GGAGGCGGAC	CAGGAGAGCT	1980
CCAGGAGGAA	AGCGCTGGTA	TGTTTATAGT	CTGTGCGCTT	TTGGGCACTT	CTGACTTGAG	2040
GGTGATTTT	TGTTGCTCTC	GTCAAGGCGG	CGGAGGCTAT	GGAAAAGCG	CAGCAACGCG	2100
GGCTTTTATC	GGTTGCTGCG	CTTTGCTGCG	CGTTTCTGCT	ACATTTTCTT	TGTTGCTGTA	2160
TGCTCTGATT	CTTGGAATAA	CGTATTTAGC	GGTTTGGAGT	GGCTGATAC	CGTTGCGGCG	2220
AGCTGAACCA	CCGAGCGGAG	GGATCTGCTG	AGCGAGGAGG	CGGAGAGCGG	CGTTATGCGG	2280
TAATTTCTCC	TTCGCACTCT	GTGCGGATAT	TCACAGCGCA	TATTTGCTG	ACCTCTAGTA	2340
CATCTGCTC	TGATCGCGCA	TGCTTAAGCC	AGTATACACT	CGCTATCGC	TGCTGATGTC	2400
GGTATGCTT	CGCGCGGAC	ACCGCGGAGT	ACCGGCTGAC	GGCGCTGAC	CGGCTGCTCT	2460
GCTCCCGCCA	TGCGCTTACA	GGCAAGCTGT	GGCGCTCTCC	GGGAGCTGCA	TGTGTGAGAG	2520
GTTTTACGCG	TGATCAAGCA	AACCGCGGAG	GGAGCTGCGG	TAGAGCTCAT	CAGGCTGCTC	2580
GTGAAGCGAT	TCAGAGATGT	CTGCTGTCTC	ATCGGCTGCG	AGCTGCTTGA	GTCTTCTCGAG	2640
AAGGCTTAT	GTGCTGCTTC	TGADAAAGCG	GGCTATGTTA	AGGCGGCTTT	TTTCTCTCTT	2700
GGTCACTGAT	GGCTCGCTGT	AAGCGGATTT	TGCTTCTCTG	GGGTAATGTA	TACCGATGTA	2760
ACGAGAGAGG	ATGCTGAGCA	TAGCGCTTAC	TGATGATGAA	CATGCGCGCT	TACTTGAAGG	2820
TTGTCAAGAT	AAACAACTCG	CGGTATGAGT	GGGCGGCGAC	CAGATAAAAG	TGACTCAAGG	2880
TCAATGCCAG	CGCTTCGTTA	ATACAGATGT	AGTGTCTTCA	CAGGCTGAGG	AGCAGCATCC	2940
TGCGATGCGG	ATCGGAGACA	TAAATGTGCA	GGCGCTGAC	TTTCTCTCTT	CGAGCTTTA	3000

CGAAGCAGGG	AAACCGAAGA	CGATTCATGT	TGTTCCTCAG	GTCCGAGACG	TTTTCGAGCA	3060
GCAGTGGGTT	CAGTTGGGCT	GGGTAAGCG	TGATTCATTC	TGCTAACGAG	TAGGACAAAC	3120
CGCCACGCTT	AGCGGGGTCC	TCAACGACAG	GAGCAGGATC	ATGCGCAGCC	GTGGGGGGCG	3180
CATCGCGGCG	ATAATGGGCT	GCTTCTGGCC	GAAAGCTTTC	GTGCGGGGAC	CAGTGAAGAA	3240
GGTTTGAGCG	AGGCGGTGCA	AGATTCCGAA	TGACGCAAGC	GACAGCGCGA	TGATPLGTGC	3300
GCTCAGCGCA	AGGCGGTGCT	CTCCGAAAT	GACCCAGAGC	GCTGCGCGCA	CGTGTCTAC	3360
GAGTTGCATG	ATAAAGAAGA	CAGTCATTAAG	TGCGGCGAGC	ATAGTCTATG	CCCGCGGCCA	3420
CGGAAAGGAG	CTGACTGGGT	TGAAGGCTCT	CGAGCGCATC	GGTCGAGATC	CGGTTGCCCA	3480
ATGAGTGAAG	TAACTTACCT	TAAATGGGTT	GCGCTCACTG	CGCGCTTTCC	AGTCGGGAAA	3540
GCTGCTGGCG	CAGCTGCATT	AATGAATGCG	CGAGCGCGCG	GGGAGAGGCG	GTTCGGGTAT	3600
TGGCGGCGAG	GCTGGTTTTT	CTTTTCACCA	GTGAGACCGG	CAGAGGTGTA	TTCGCTTTCA	3660
CGGCTTGCGT	CTGAGAGAGT	TGCGAGCAGC	GCTCCAGGCT	GCTTTGGGCC	AGCAGGGGAA	3720
AATCTCTTTT	GATGTGGGTT	AAGGGCGGGA	TATAACATGA	GCTGTCTTCG	GPATCTCTGT	3780
ATCCCACTAC	CGAGATATCC	GCACCAACGC	GCAGCGCGGA	CTCGGTAATG	CGCGGCGTTG	3840
CGCGCAGCGC	CATCTGATCG	TGGCAACCCA	GCATCGCAGT	GGGACGAGTG	CCCTGATTCG	3900
GCATTTGCAT	GCTTTGTGTA	AAACCGGACG	TGCGACTGCA	GTGCGCTTCG	GCTTCGGGTA	3960
TGGGCTGAAT	TTGATGGGCA	GTGAGATATT	TATGCCAGGC	AGCGAGAGCG	AGACCGCGCG	4020
AGACAGAACT	TAATGGGCGC	GCTAACAGCG	CGATTTGGTG	GTGACCGCAAT	CGGACGAGAT	4080
GCTCCAGCGC	CAGTGCAGTA	CGTCTTGCAT	GGGAGAAAT	AATAGCTTTG	ATGGGTGCTT	4140
GCTCAGAGAC	ATCAGCAAAAT	AACGCGGAAA	CATTATGCGA	GGCAGCTTCC	ACAGCAATGG	4200
CATCTCTGTC	ATCCAGCGGA	TAGTTAATGA	TGAGCGCGCT	GAGCGGTTGC	CGGGAAGGAT	4260
TGTCCAGCGC	CGCTTACAGC	GCTTCGAGCG	CGCTTCGTTG	TAGCATTCAG	ATCAGCAGCG	4320
TGGCAGCGAG	TGATGCGGCG	CGAGTTTAA	TGCGCGGACG	AATTTGGGAG	CGCGGCTGCA	4380
GGGCGAGAGT	GGAGCTGGCA	ACGCGCAATCA	CGAGCGAGTG	TTGCGCGGCG	AGTTGTTGTC	4440
CGAGCGGCTT	GGAGATTTAA	TTCAGCTTCG	CGATCGCGCG	TTCCACTTTT	TGCGCGGTTT	4500
TGCGAGAAAC	GTCTCTGGCC	TGCTTACGCA	CGCGGAGAGC	GCTCTGATAA	GAGACAGCGG	4560
CATCTCTGCG	GACATCGGAT	AACCTTATCT	GTTCACATTT	CACGACGCTG	AATTAAGCTCT	4620
CTTCCGCGCG	CTATGATGCG	ATACCGCGAA	AGGTTTTGCG	CGATTCGATG	GTCTCGGCGA	4680
TCTCGAGGCT	CTCTCTTATG	CGACTCTCTC	ATTAGGAGCG	AGCGCAGTAG	TAGGTTGAGG	4740
CGCTTGAGCA	CGCGCGCGCG	AGGGAATGCT	GATTCGAGCG	AGATGGCGCG	CGACATCTCC	4800
CGGCGCAGCG	GGCTGCGCAC	CATACCGAGG	CGGAGACAGG	CGCTCATGAG	CGCGAAGTGG	4860

CGAGCCCGAT	CTTCCCATX	GCTGATGTCG	GCGATATAGG	CGCCAGCAAC	CGCAGCGTG	4928
GGCCCGATGA	TGGCGGCCAC	GATGCGTCCG	GCGTAGAGCA	TGGAGATCTC	GATCCCGCGG	4980
AAITAAATACG	ACTCACTATA	GCGGAAITGT	GAGCGGATAA	CAATTCGCTT	CTAGAAATAA	5040
TTTTCTTTAA	CITTAAGAAG	GAGATATACA	TATCGGCGAT	CATCATCATC	ATCAGCTGAT	5100
CGCATCATTC	GCGACGAGCC	CCACATATTC	GGAACAGCGC	GCGCGGAGG	CGGTCCAGCG	5160
GGCGCGGGAT	AGCGTCGATC	ACATCCGCGT	GCTCGGCTC	ATTGAGCGAG	ACATGCGCGT	5220
GGACAGCGCC	GGCAAGATCA	CCATCCGAT	CAAGCTCGAA	GTGTCGTCAC	AGATGAGGCC	5280
GGCGCAACCG	AGCGGCTGAA	AGCGCGGAG	GCTTCGCTC	GAAACGCGG	CGCGCCCGCG	5340
TACTGTCGCG	ACTACCCCGG	CGTCTCGG	GCTGACCTTC	GCGGAGCGCG	GTACGAGCTT	5400
GCTCTACCGG	CTGTCGACCC	TGTCGGGCGC	GCGCTTTCAC	GAGAGCTATC	CGAGCGCTAC	5460
GATCACCGCT	CAGGCGGCGC	GTTCTGCTTC	CGGATCGCG	CAGGCGCGCG	CGGAGCGCTT	5520
CAACATTCGG	GCTTCGAGCG	CTATCTCTC	GGAAGGATGAT	ATGCGCGCGG	ACAAGCGCGT	5580
GATGAAGATC	GCGCTAGCCA	TCTCGGCTCA	CGAGCTCGAC	TAGAACCTGC	CGGAGCTGAG	5640
CGAGCACCTC	AGGCTGAAGC	GAAAGTCTCT	GCGGCGCTTC	TAGCAGGCGA	CGATCAAAAC	5700
CTGCGAGCAC	CGCGAGATCG	CTGCGCTCAA	CGCGCGGCTG	AACTGCGCGG	CGACCGCGCT	5760
AGTTCGCGTG	CAGCGCTCGG	AGCGCTCGG	TAGACCTTTC	TGCTGACCG	AGTACGCTTC	5820
CAGCAAGAT	CGCGAGGCTT	GCGCGAGCTC	GCGCGCTTC	GCGAGCGCGG	TGAGCTTCCT	5880
GCGCGTTCGG	GCTGCGGCTG	GTCAGAGCGG	CAAGCGCGCG	ATGCTGAGCG	GTCGCGCGGA	5940
GACACCGAGC	TGCTGCGCT	ATACCGCAT	CAGCTTCCTC	GCGCGAGCGA	CTCAGCGCGG	6000
ACTGCGCGAG	GCTCAGCTAC	GATATAGCTC	TGCGATTTTC	TTCTTCTCGG	AGCGCGAAG	6060
CATTCAAGCG	GCGCGGCTG	GCTTCGCTTC	GAAACGCGG	GCGAGCGAG	CGATTTCTAT	6120
GATCGAGCGG	CGCGCGCGCG	AGCGCTAGCG	GATCATCAAC	TAGAGCTAGG	CGATCTCTAA	6180
CGACCGCGAA	AGAGCGCGCG	CGCGCGCGGA	GAGCTTGAAG	GCTTTCTCTG	ACTGCGCGAT	6240
CAGCGAGCGC	ACAAGCGCTT	GTTCTCTCGA	CGAGCTTCAT	TTCGCGCGAC	TGCGCGCGCG	6300
GCTGCTGAAG	TCTCTCTAGG	CGTCTCTCTG	GAGCTTTCTG	AGCGCTGAGA	TGAAGCGCGA	6360
TGCGCTTACG	CTGCGCGAGG	AGCGAGCTAA	TTCGAGCGCG	ATCGCGCGCG	ACTGGAAGAC	6420
CGAGCTGAC	CAGCTGAGAT	CGAGCGCGAG	TTCCTTCTAG	GCGAGCTGCG	GCGCGCGCGC	6480
AGCGAGCGCG	GCGCGAGCGG	GCGCTGCTCG	CTTCGAAGAA	CGAGCGAGAA	AGCGAGCGCA	6540
GGAGCTCGAG	GAGATCTCGA	CGATATCTCG	TGCGCGCGCG	GTCGATATCT	CGAGCGCGCGA	6600
CGAGAGCGAG	CAGCAGCGCG	TGCTCTCGCA	AGCGCGCTTT	GTCGCGAGAA	CGCGCGCGCTC	6660
GCGCGCTCG	ACTGCTGCGG	GCGCGCTCGG	AGCGCGCGCG	GCTGCTGCGG	CGCGCGCGCTC	6720



```

GGCGCGCGCC AACACGTCGA ATGCCAGCCG GGGGATGCC AACCGAGAAC CTCGCGCGCG 6780
CGACCCGAAAC GCACCGCCGC CACCTGTCTT TCGCCCAAC GCACCGCAAC CAGTCGCGAT 6840
CGACAACCCG GTTGGAGGAT TCAGCTTCGC GTTGGCTGCT GGTGCGGTGG AGTCTGACGC 6900
CGCGCACTTC GACTACGGTT CAGCAGTCCG CAGCAAAAGC ACCGGGAGCC CGCCATTTC 6960
CGGAGAGCGG AXXCGGTGG GAAATGACAC CGTATCTGTC CTGGGCGGC TAGACCAAAA 7020
GCTTTACGCC AGCGCGAAT CCACTGACTG CAGCGCGCG GCGCGTTGG GCTCGGACAT 7080
GGTTHAATTC TATATGCCCT ACGCGGACAC CGGATCAAC CAGGAACCG TCTCGCTTGA 7140
CGCCACGCGG GTGCTGGA ACGCGTCTTA TTACGACTC AACTCGAGCG ATCCGAGTAA 7200
CGCGACGCGC CAGATCTGGA CCGGGTAAAT CGCTCGGCC GCGGCGAAG CACCGGCGCC 7260
CGGCGCCCTT CAGCGCTGCT TTGTTGATG GCTCGGACC GCAACAAC CCCTGGACAA 7320
GGGCGCGGCC AAGGCGCTGG CGCAATCTAT CCGCGCTTTC CTGCGCGCG CCGCGGCGCC 7380
GGACCGGCTT CTTGAGAGC CGACTCGGCG GCGGCGCGG GCGGCGAAG TCGCTCTTAC 7440
CCGACGACA CCGACACCGC ACGCGACCTT ACGGCGTGA GATTCTGCG CATATCATC 7500
ACACTGCGCG CCGCTGAGG ACCACCGCA CACGCTGA GATCGCGCT CTACCAAGC 7560
CGCAAGGAA GCTGAGTGG CTCTGCGAC CGCTGAGCA TACTATCAT AACCGCTGG 7620
GGCTCTAAA CGCTCTTGA GCGTTTTT GCTCAAGGA GCACTATAT CGGAT 7676

```

(2) INFORMATION FOR SEQ ID NO:214:

- (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 802 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

```

Met Gly His His His His His Val Ile Asp Ile Ile Gly Thr Ser
  1             5             10             15
Pro Thr Ser Trp Glu Gln Ala Ala Ala Glu Ala Val Gln Arg Ala Arg
  20             25             30
Asp Ser Val Asp Asp Ile Arg Val Ala Arg Val Ile Gln Gln Asp Met
  35             40             45
Ala Val Asp Ser Ala Gly Lys Ile Thr Tyr Arg Ile Lys Leu Glu Val
  50             55             60
Ser Phe Lys Met Arg Pro Ala Gln Pro Arg Gly Ser Lys Pro Pro Ser
  65             70             75             80

```

Gly Ser Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro  
 85 90 95  
 Ala Ser Ser Pro Val Thr Leu Ala Glu Val Gly Ser Thr Leu Leu Tyr  
 100 105 110  
 Pro Leu Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn  
 115 120 125  
 Val Thr Ile Thr Ala Gln Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln  
 130 135 140  
 Ala Ala Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser  
 145 150 155 160  
 Glu Gly Asp Met Ala Ala His Lys Gly Asn Met Asn Ile Ala Leu Ala  
 165 170 175  
 Ile Ser Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His  
 180 185 190  
 Leu Lys Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile  
 195 200 205  
 Lys Thr Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn  
 210 215 220  
 Leu Pro Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly  
 225 230 235 240  
 Asp Thr Phe Leu Phe Thr Gln Tyr Leu Ser Lys Glu Asp Pro Glu Gly  
 245 250 255  
 Trp Gly Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val  
 260 265 270  
 Pro Gly Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys  
 275 280 285  
 Ala Gln Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp  
 290 295 300  
 Gln Ala Ser Gln Arg Gly Leu Gly Glu Ala Gln Leu Gly Asn Ser Ser  
 305 310 315 320  
 Gly Asn Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala  
 325 330 335  
 Gly Phe Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp  
 340 345 350  
 Gly Pro Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile  
 355 360 365  
 Val Asn Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala  
 370 375 380  
 Phe Leu His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp  
 385 390 395 400  
 Glu Val His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp  
 405 410 415

Ala Leu Ile Ala Thr Ile Ser Ser Ala Glu Met Lys Thr Asp Ala Ala  
 420 425 430  
 Thr Leu Ala Glu Gln Ala Gly Asn Phe Glu Arg Ile Ser Gly Asp Leu  
 435 440 445  
 Lys Thr Glu Ile Asp Gln Val Glu Ser Thr Ala Gly Ser Leu Gln Gly  
 450 455 460  
 Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala Ala Val Val Arg  
 465 470 475 480  
 Phe Gln Glu Ala Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu Ile Ser  
 485 490 495  
 Thr Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg Ala Asp Gln Glu  
 500 505 510  
 Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe Val Pro Thr Thr Ala  
 515 520  
 Ala Ser Pro Pro Ser Thr Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro  
 530 535 540  
 Val Ala Pro Pro Pro Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro  
 545 550 555 560  
 Gly Asp Pro Asn Ala Ala Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro  
 565 570 575  
 Pro Pro Val Ile Ala Pro Asn Ala Pro Gln Pro Val Arg Ile Asp Asn  
 580 585  
 Pro Val Gly Gly Phe Ser Phe Ala Leu Pro Ala Gly Trp Val Glu Ser  
 590 600  
 Asp Ala Ala His Phe Asp Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr  
 610 615 620  
 Gly Asp Pro Pro Phe Pro Gly Gln Pro Pro Pro Val Ala Asn Asp Thr  
 625 630 635 640  
 Arg Ile Val Leu Gly Arg Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu  
 645 650 655  
 Ala Thr Asp Ser Lys Ala Ala Ala Arg Leu Gly Ser Asp Met Gly Gln  
 660 665 670  
 Phe Tyr Met Pro Tyr Pro Gly Thr Ala Ile Asn Gln Glu Thr Val Ser  
 675 680 685  
 Leu Asp Ala Asn Gly Val Ser Gly Ser Ala Ser Tyr Tyr Gln Val Lys  
 690 695 700  
 Phe Ser Asp Pro Ser Lys Pro Asn Gly Gln Ile Trp Thr Gly Val Ile  
 705 710 715 720  
 Gly Ser Pro Ala Ala Asn Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp  
 725 730 735  
 Phe Val Val Trp Leu Gly Thr Ala Asn Asn Pro Val Asp Lys Gly Ala

211

740	745	750
Ala Lys Ala Leu Ala Glu Ser Ile Arg Pro Leu Val Ala Pro Pro Pro		
755	760	765
Ala Pro Ala Pro Ala Pro Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala		
770	775	780
Gly Glu Val Ala Pro Thr Pro Thr Thr Pro Thr Pro Glu Arg Thr Leu		
785	790	795
Pro Ala		800

## CLAIMS

1. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 120)
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID No. 121)
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 122)
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID No. 123)
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID No. 124)
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 125)
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID No. 126)
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID No. 127)
- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID No. 128) and
- (j) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID No. 136)

wherein Xaa may be any amino acid.

2. A polypeptide comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative

substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID No. 129) and
- (b) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No. 137), wherein Xaa may be any amino acid.

3. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101 or a complement thereof under moderately stringent conditions.

4. A polypeptide comprising an immunogenic portion of a *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 26-51, 138, 139, 163-183 and 201, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 26-51, 138, 139, 163-183 and 201 or a complement thereof under moderately stringent conditions.

5. A DNA molecule comprising a nucleotide sequence encoding a polypeptide according to any one of claims 1-4.

6. An expression vector comprising a DNA molecule according to claim 5.

7. A host cell transformed with an expression vector according to claim 6.

8. The host cell of claim 7 wherein the host cell is selected from the group consisting of *E. coli*, yeast and mammalian cells.

9. A pharmaceutical composition comprising one or more polypeptides according to any one of claims 1-4 and a physiologically acceptable carrier.

10. A pharmaceutical composition comprising one or more DNA molecules according to claim 5 and a physiologically acceptable carrier.

11. A pharmaceutical composition comprising one or more DNA sequences recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and a physiologically acceptable carrier.

12. A vaccine comprising one or more polypeptides according to any one of claims 1-4 and a non-specific immune response enhancer.

13. A vaccine comprising:  
a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and  
a non-specific immune response enhancer.

14. A vaccine comprising:  
one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140 and 141, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and  
a non-specific immune response enhancer.

15. The vaccine of claims 12-14 wherein the non-specific immune response enhancer is an adjuvant.

16. A vaccine comprising one or more DNA molecules according to claim 5 and a non-specific immune response enhancer.

17. A vaccine comprising one or more DNA sequences recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and a non-specific immune response enhancer.

18. The vaccine of claims 16 or 17 wherein the non-specific immune response enhancer is an adjuvant.

19. A pharmaceutical composition according to any one of claims 9-11, for use in the manufacture of a medicament for inducing protective immunity in a patient.

20. A vaccine according to any one of claims 12-18, for use in the manufacture of a medicament for inducing protective immunity in a patient.

21. A fusion protein comprising two or more polypeptides according to any one of claims 1-4.

22. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and ESAT-6.

23. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and the *M. tuberculosis* antigen 38 kD (SEQ ID NO:155).

24. A pharmaceutical composition comprising a fusion protein according to any one of claims 21-23 and a physiologically acceptable carrier.

25. A vaccine comprising a fusion protein according to any one of claims 21-23 and a non-specific immune response enhancer.

26. The vaccine of claim 25 wherein the non-specific immune response enhancer is an adjuvant.



27. A pharmaceutical composition according to claim 24, for use in the manufacture of a medicament for inducing protective immunity in a patient.

28. A vaccine according to claims 25 or 26, for use in the manufacture of a medicament for inducing protective immunity in a patient.

29. A method for detecting tuberculosis in a patient, comprising:

(a) contacting dermal cells of a patient with one or more polypeptides according to any one of claims 1-4; and

(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

30. A method for detecting tuberculosis in a patient, comprising:

(a) contacting dermal cells of a patient with a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and

(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

31. A method for detecting tuberculosis in a patient, comprising:

(a) contacting dermal cells of a patient with one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203; and

(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

32. The method of any one of claims 29-31 wherein the immune response is induration.

33. A diagnostic kit comprising:
- (a) a polypeptide according to any one of claims 1-4; and
  - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
34. A diagnostic kit comprising:
- (a) a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
  - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
35. A diagnostic kit comprising:
- (a) a polypeptide encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203; and
  - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
36. A diagnostic kit comprising:
- (a) a fusion protein according to any one of claims 21-23; and
  - (b) apparatus sufficient to contact said fusion protein with the dermal cells of a patient.
37. A fusion protein according to claim 23 comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 153 and 209.

1/11

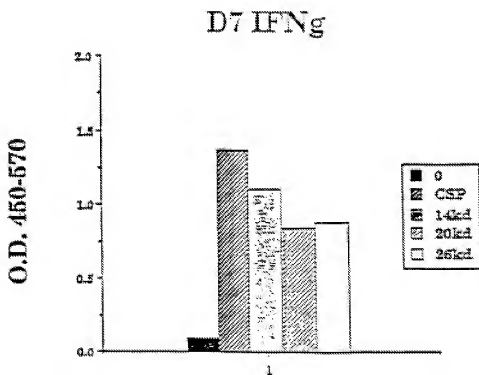
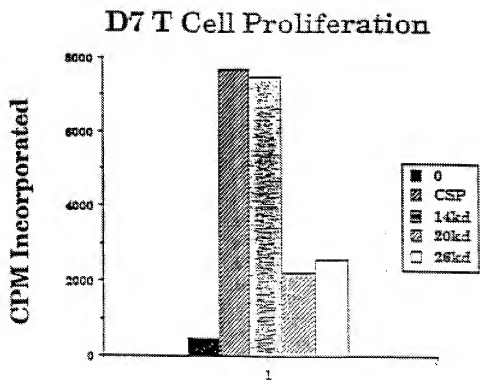
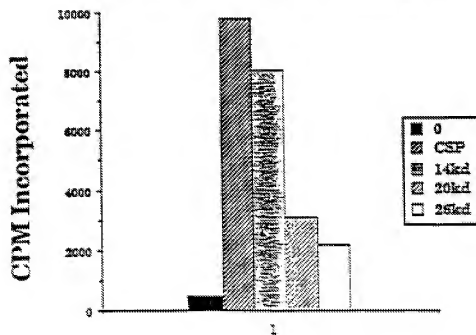


FIG. 1A

2/11

# D160 T Cell Proliferation



# D160 IFN $\gamma$

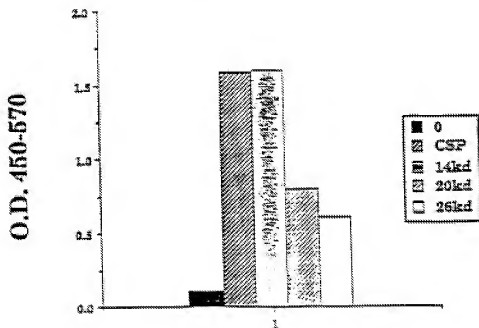


FIG. 1B

3/11

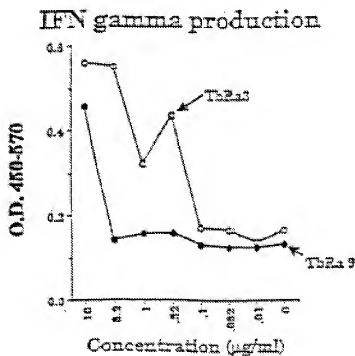
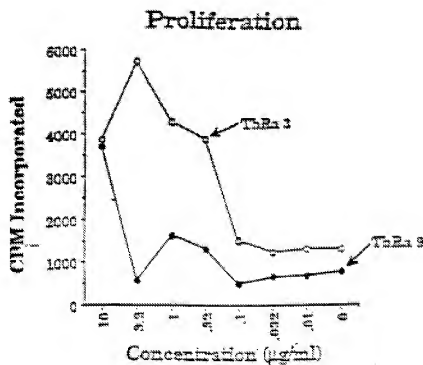
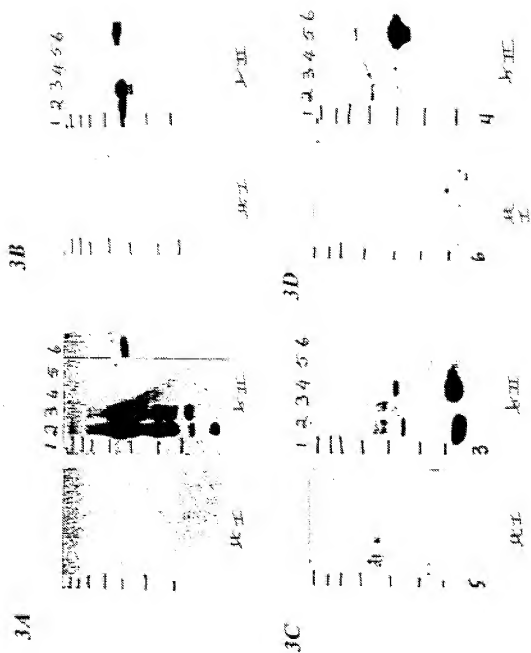


FIG. 2

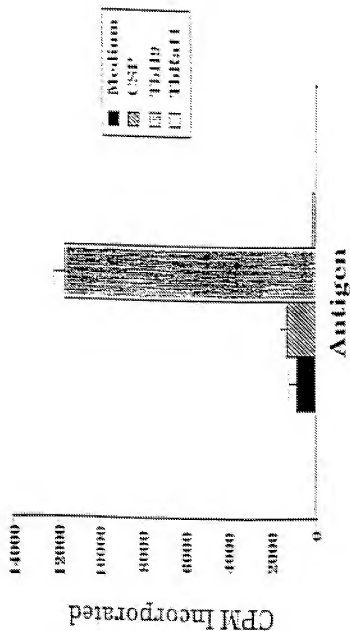
4/11



FIGS. 3A-D

5/11

# T cell clone 131TbH9 responds poorly to CSP



6/11

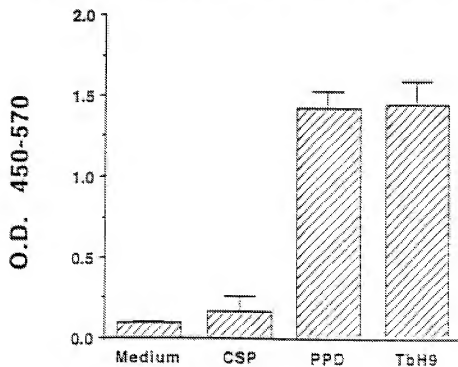
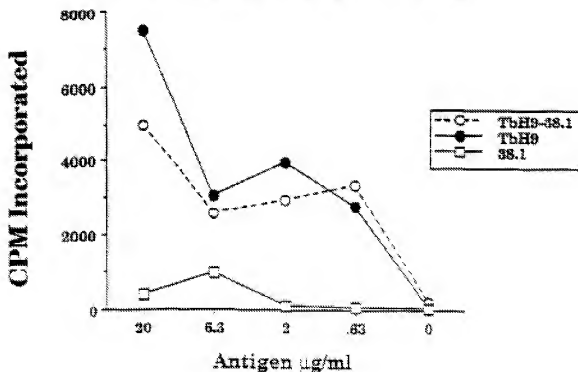
T Cell Clone PPD 800-10 IFN $\gamma$  Production

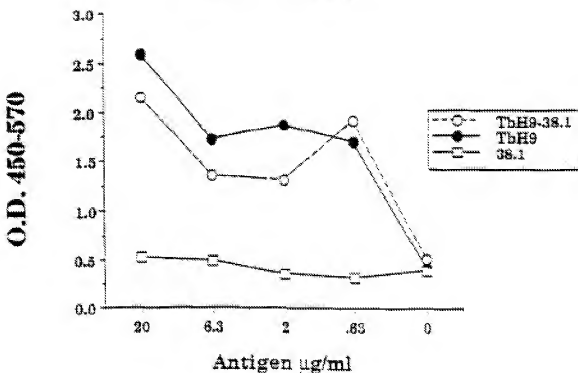
FIG. 4B



# D131 T Cell Proliferation



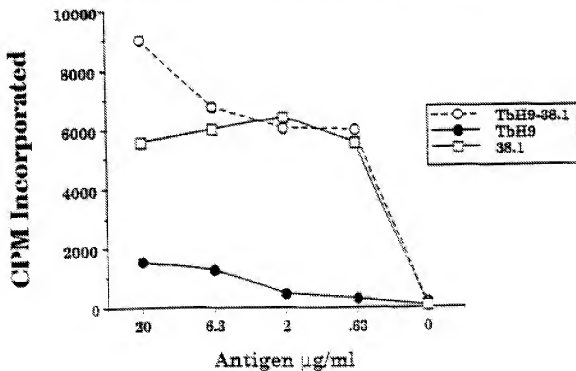
## D131 IFN $\gamma$



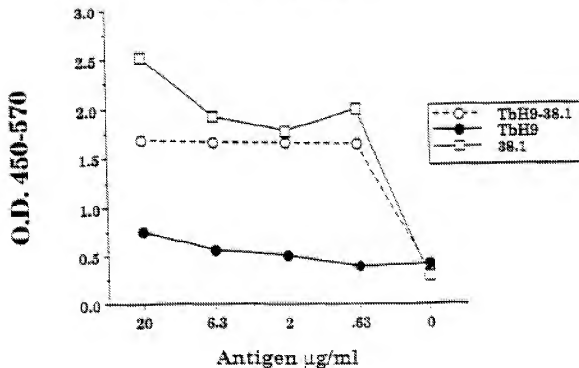
FIGS. 5 A-B

8/11

# D184 T Cell Proliferation

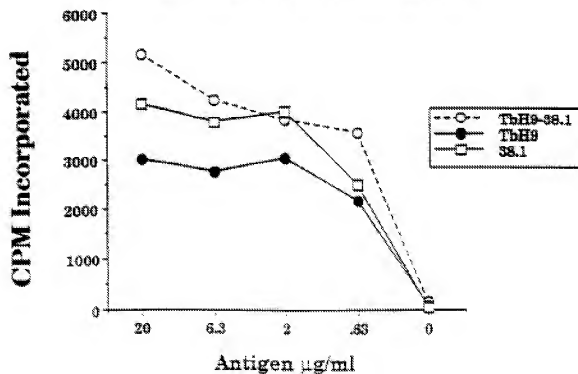


## D184 IFN $\gamma$

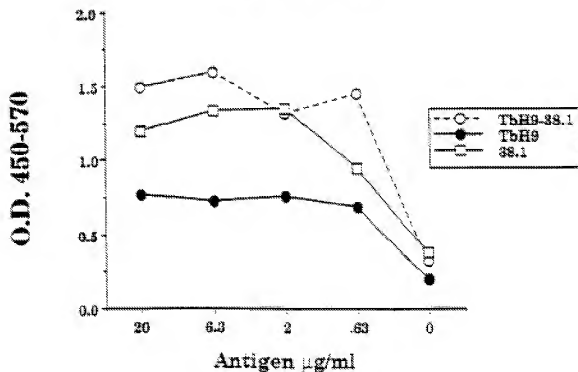


FIGS. 6 A-B

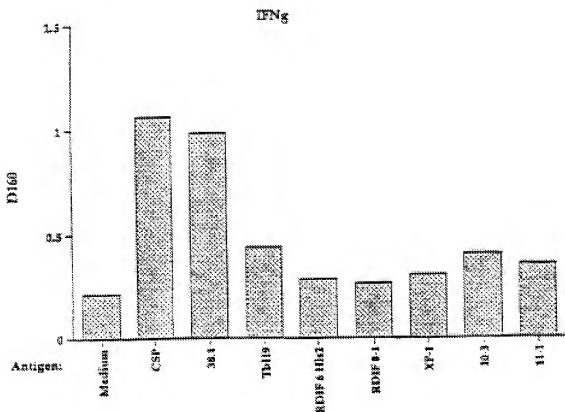
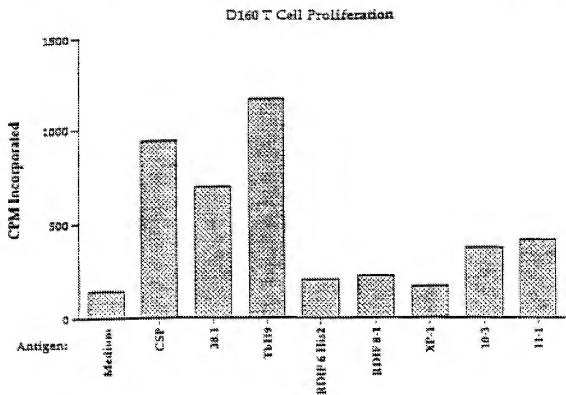
## D201 T Cell Proliferation

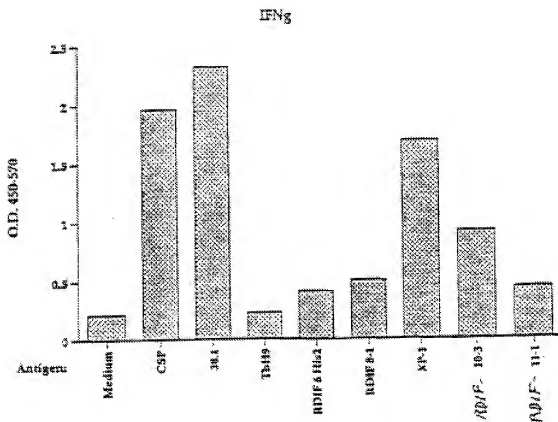
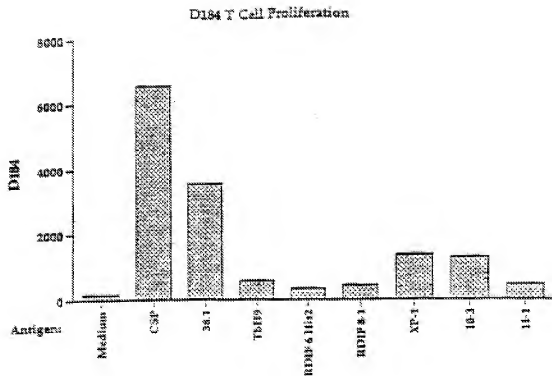


## D201 IFN $\gamma$



FIGS. 7 A-B

**FIGS. 8A-B**

**FIGS. 9A-B**